JOURNAL OF APPLIED PHYSICS

JOURNAL OF APPLIED

PHYSICS

(Formerly PHYSICS)

MARCH, 1938 Vol. 9, No. 3

In This Issue

Page

Relation of Physics to the Biological Sciences DETLEV W. BRONK 139

Electron Diffraction Methods of Studying Organic Films L. H. GERMER 143

Biophysics and Biochemistry of Viruses W. M. STANLEY 148

Photochemical Basis of Vision SELIG HECHT 156

Mechanics of Muscular Contraction in Man W. O. FENN 165

Some Aspects of Nuclear Physics of Possible Interest in Biological Work L. A. DUBRIDGE 178

Contributed Original Research

Quantitative Measurements of Texture by the Magnetic Torque Method L. P. TARASOV 192

Invariant Form of the Maxwell-Lorentz Equations for Accelerated Systems GABRIEL KRON 196

Time Constants of Incandescent Lamps W. E. FORSYTHE, M. A. EASLEY AND D. D. HINMAN 209

In Every Issue

Here and There	187
Innovations in Instruments	188
Some of the New Books	189
Contributors to this Issue	191
Current Literature of Physics	vi
Advertising Index	xx

BOARD OF EDITORS

ELMER HUTCHISSON, *Editor*

Associate Editors

R. M. BOZORTH
D. W. BRONK
W. P. DAVEY
O. S. DUFFENDACK
SAUL DUSHMAN
A. STUART HUNTER
C. H. KUNSMAN
J. T. LITTLETON
L. L. NETTLETON
W. B. NOTTINGHAM
A. R. OLPIN
JOHN T. TATE

Publication of the American Institute of Physics

ADMINISTRATIVE STAFF

HENRY A. BARTON, *Director*

MADELINE M. MITCHELL, *Publications Manager*

WILLIAM C. COPP, *Advertising Manager*

GOVERNING BOARD

JOHN T. TATE, *Chairman*
KARL T. COMPTON
E. C. CRITTENDEN
WHEELER P. DAVEY
HOMER L. DODGE
HARVEY FLETCHER
ARTHUR C. HARDY
VERN O. KNUDSEN
MELVIN MOONEY
FREDERIC PALMER, JR.
R. L. PEEK, JR.
GEORGE B. PEGRAM
W. B. RAYTON
F. K. RICHTMYER
WALLACE WATERFALL

THE JOURNAL OF APPLIED PHYSICS, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as PHYSICS. It aims to be of service not only in physical laboratories, but also in the laboratories of industrial, chemical, geological, geophysical, meteorological, radio, and similar concerns. THE JOURNAL OF APPLIED PHYSICS publishes editorials, news of physicists and reviews, as well as technical papers of applied physics.

SUBSCRIPTION PRICE

	U. S., POSS. AND CANADA	FOREIGN
To members of scientific societies	$5.00	$5.70
To all others	7.00	7.70

Manuscripts should be submitted to Elmer Hutchisson, Editor, University of Pittsburgh, Pittsburgh, Pennsylvania.

Changes of address, proofs, and complaints of failure to receive the JOURNAL OF APPLIED PHYSICS should be addressed to the Publications Manager. New copies can be sent free only if notice of nondelivery is received within three months of date of issue.

Subscriptions and orders for back numbers should be addressed to Prince and Lemon Streets, Lancaster, Pa. or to the American Institute of Physics, 175 Fifth Avenue, New York, N. Y.

Advertising rates supplied on request. Orders, advertising copy and cuts should be sent to the American Institute of Physics.

Entered as second class matter at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1897. Accepted for mailing at the special rate of postage provided for in the Act of February 28, 1925, authorized May 2, 1932.

Copyright 1938, by the American Institute of Physics.

Journal of Applied

Physics

Volume 9 March, 1938 Number 3

The Relation of Physics to the Biological Sciences*

BY DETLEV W. BRONK

Eldridge Reeves Johnson Foundation
University of Pennsylvania, Philadelphia

I

THE widening frontiers of modern science and the multiplication of experimental methods have brought about severe specialization. This specialization has been justified by its accomplishments, for without it a worker is hampered by superficial and inadequate understanding of the facts which immediately surround the problem he is investigating. The time has now come, however, when the practical development of science requires individuals and agencies devoted to the correlation of its subdivisions. A synthesis of the sciences is likewise necessary if they are to retain their broader cultural values, for only thus do we see the detailed facts in their true perspective as parts of a more complete pattern of knowledge. This dual need for greater specialization and more integration in science is the problem and the opportunity of the biophysicist. It is the need which has prompted this symposium.

II

There are wide differences of opinion as to the desirability of establishing biophysics as a new subdivision of organized science. Without arguing this question we must admit that it is a useful term which has come to designate a certain field of scientific endeavor. It is, therefore, desirable, at the start of this symposium, to consider what is its scope. There are many who believe it to be the development of physical instruments and methods for biological and medical work, but this is no more important a part of biophysics than instrument design is

* Presented at the Symposium on Biophysics, Philadelphia, Pennsylvania, November 4, 5 and 6, 1937.

of any other branch of science. The true province of biophysics is revealed by a consideration of the fact that a large part of biological research is devoted to the analysis and explanation of living mechanisms in terms of physical concepts and principles; that much of applied biology is concerned with the control of vital processes by physical agents.

After careful thought I can recognize but three basic disciplines in biology, and these are anatomy, biochemistry and biophysics, the last two constituting the inclusive science of physiology. On first thought, it may appear necessary to add another category to include problems of organization, for they represent one of the most important phases of biological study. One can readily show, however, that all of those problems ultimately depend on the anatomical arrangements of the body and upon the physical and chemical processes which control the association of its several parts. Finally, these three basic divisions of biology are themselves so intimately related that we can with difficulty define their boundaries. Such inseparability of biochemistry and biophysics will be readily apparent to this audience, but it may be well to emphasize the essential role of the physical sciences in anatomy.

It is a fundamental principle of biology that the functional properties of a living system are determined by its structure. It is true of the gross structures which can be observed and examined with the naked eye, for they determine the nature of the interrelation between the organs and between groups of cells. It is also true of the finer structure which we see with the microscope, for that reveals the character of intercellular relationships and is often an index of cell structure. That cellular structure, in turn, is determined by the submicroscopic anatomy of the cell which must ultimately be explained in terms of molecular arrangements and configurations, and is, therefore, a problem in physical chemistry. It is this structure which really determines the behavior of the cell, and when we analyze living systems at this level, the division between structure and function disappears, and biology as a whole becomes more nearly a branch of physics and chemistry.

If we consider the fundamental structure of a cell in these terms, we find that it is not a fixed, static entity such as that which we view in a histological preparation. It is, as I have indicated, an assemblage of molecules and atoms held together in a certain pattern by electrical forces. This pattern is not stable. It is continually tending to disintegrate, and only by an expenditure of energy can it maintain itself in that state of dynamic equilibrium which characterizes a living cell. Any alteration in its physical and chemical environment throws this unstable system out of balance, and thereupon a wave of change spreads along the nerve fiber as an impulse, or the tension developed by a muscle fiber increases, or a gland cell discharges its secretion. After the intimate structure of the cell has been thus disorganized, additional energy is required to reform the original state. It therefore follows that a cell is continually requiring energy in order to maintain itself in such a condition that it will be able to go into action, and, having been in action, still more energy is necessary to prepare it for further activity.

It is apparent, therefore, that functional biology and medicine are largely concerned with the energetics of living systems, and that is the scope of biophysics. It is, accordingly, concerned with the manner in which the organized living state is maintained by an expenditure of energy; with the functional effects induced by alterations in the energy supplied to the system; and with the energy changes in a cell or tissue which are an index of its activity, and which reveal the nature of the living mechanisms—as, for instance, the heat produced by nerve and muscle or the bielectric action potential.

III

Because the biological and physical sciences are so intimately related, much of the research and advanced teaching in biology depends upon the competent use of physical methods and principles. How this can be accomplished is a problem of considerable importance. The most generally advocated solution is the cooperation between physicists and biologists. This may be effective when the physicist is merely called upon to devise an experimental method, or when the attack upon a problem can be resolved into several approaches, some of which are primarily physical, and others biological. But even then

such cooperative work suffers one serious disadvantage. Biological research must generally be conceived and directed by a well informed biologist, and only he can interpret the results and formulate the generalizations. The cooperating physicist is, therefore, seldom more than a technical assistant, and it is not surprising that young physicists of real ability will not undertake such work.

Among the organized divisions of science biophysics is more a part of biology than of physics. It is true that an understanding of biological mechanisms can only be attained in terms of physical and chemical principles and concepts, but there is a large body of facts and theories concerning living organisms which is unknown to the physicist. Without that knowledge and the point of view which it engenders, even the best trained physicist will be unable to formulate, direct and interpret biophysical problems. Effective work in this field, consequently, requires a thorough training in both biology and physics. The formulation of the more important integrated generalizations in the biological sciences now requires individuals who can draw upon a wide range of sources, critically and with understanding.

IV

In order to achieve this end a certain number of men originally trained as physicists have taken a rigorous course of discipline in the biological sciences; some biologists have struggled to qualify as competent physicists. It is my personal conviction that a more satisfactory procedure is the simultaneous study of the physical and biological sciences throughout the graduate years. This is the plan we have followed in the Johnson Foundation, and we consider the results to be excellent. Thus the student develops the capacity to think of biological problems as a physicist, and does not create in his mind artificial barriers between the subdivisions of science which are merely imagined or which have been developed only for budgetary and administrative purposes. Scientists with this breadth of training and viewpoint will certainly play a major role in the biology of the future.

It is, however, difficult to find many students qualified either by interest or undergraduate training to undertake such a course of graduate study and research. This is due to the fact that college students generally specialize in either the physical or the biological sciences and seldom do extensive work in both fields. In part this is a consequence of archaic and inflexible arrangements of curricula. In larger part it is due to a still prevalent notion that an individual has either a descriptive or an analytical type of mind, and is, therefore, qualified to undertake only one or the other of the sciences. Personally, I doubt whether there is any basis for such a classification of mental powers. Be that as it may, the biological sciences, now grown out of their descriptive phase, demand of their devotees the same mental characteristics as do the physical sciences. It is, therefore, important at the present time to set up combined curricula in the biological and physical sciences for the undergraduate who looks forward to a fruitful career in modern biology or medicine.

V

This correlation of the physical and biological sciences in undergraduate schools would be greatly aided by the appointment of biophysicists to the departments of physics in our larger universities. I have had little difficulty persuading departmental chairmen that this is desirable, but its accomplishment is generally blocked for the reason that the same budgetary appropriation for the salary of a spectroscopist, a nuclear physicist, or a cosmologist brings the department more glory among physicists. Of this I have no doubt, but I still believe that the time will come when every large department of physics will afford at least one member who has been trained as a biophysicist, and is working on problems in the borderland between physics and biology.

Such an individual would have at least three important functions. He would be an admirable director of the course or curriculum in physics for premedical students and for those with a primary interest in biology. I hasten to add that I advocate no specialized or emasculated course in physics for such students. As a former teacher of physics, of physiology, and of zoology, I would say that the student needs a more rigorous and thorough training in the fundamentals—not

less—in order that he may develop greater versatility and resourcefulness for the complicated problems of biophysics. But a familiarity with the nature and problems of biology would enable the teacher of physics, who is also a biophysicist, to employ illustrations drawn occasionally from biology, as is now done from engineering. He would be able to develop in the student's mind the conception that physical laws apply to the atoms and molecules of living systems as well as to those of inorganic matter.

Such a man would bring to the physicist news in his own language from the frontiers of biology. This would broaden the intellectual outlook of the physicist and might suggest to him new problems in his own field or clarify old ones. For it is not unlikely that the study of matter organized as a living system and a consideration of its physical properties may be as stimulating to the physics of the future as it was in the days of Galvani and Volta. And in this connection it may be well to bear in mind that the experiments of the physical sciences are performed and observed by a living organism. The mechanisms of that organism may ultimately prove to be of great significance in the evaluation of physical data.

This biophysicist with his home in the physical laboratory would also serve as a valuable aid to the biologists and clinicians, within his university. He would be able to advise and collaborate with them in their research and could interpret to them the new discoveries of physics and its long-established principles. He would be invaluable to the many biophysicists who are located in biological and medical research laboratories and teaching departments.

The number of men holding such positions is gradually increasing. As the analysis and control of living mechanisms are carried to a more fundamental stage, research laboratories are increasingly dependent upon biophysicists. With the more general realization that biophysics is a basic part of all biology the personnel of the several departments of the natural sciences will include more men trained in both physics and biology. Thus bacteriology and immunology are already keenly interested in the problems of surface structure and molecular arrangements; medicine is concerned with problems of heat exchange and the hydrodynamics of blood-flow; geneticists must analyze the physico-chemical properties of the colloids in a chromosome; a new era in neurology depends upon the electrical signs of nervous activity. Every field of modern medicine has physical problems of the most exacting nature. They are being solved by an increasing number of zoologists, botanists, physiologists, neurologists, radiologists, and others who can also qualify as competent biophysicists.

The dissolution of unnecessary boundaries between biology and chemistry has been greatly aided by the establishment of institutes of biochemistry, and the progress of chemical biology has thus been accelerated on a sound basis. There is a similar need for institutes of biophysics in more of our large university centers. Such an institute of biophysics would be a powerful adjunct to all of the biological departments of a university. For there would be gathered together physical equipment and facilities which it would be economically impossible to install in each of the many departments of a medical school or division of biology. There, too, would be gathered together a group of physicists, engineers, physical chemists and biophysicists cooperating on problems of mutual interest and providing for each other technical aid and intellectual companionship which they would lack if they were scattered as individuals throughout the several existing departments. In such an institute, therefore, a worker in any field of biology or medicine would find experts and equipment to aid him, and there he would be provided a laboratory and facilities for carrying on those phases of his research which lead deep into the province of biophysics. There, too, the distant frontiers of this borderland science would be explored to the future advantage of biology.

Much of what I have said about the relation of physics to the biological sciences will be best illustrated by the accomplishments reported in the papers of this symposium. The most significant result of that relationship may not appear, however, to the casual reader. It is the satisfaction which comes to the workers in this field, for we have a rare opportunity to glimpse the essential unity of science. To comprehend this is the final objective of every natural philosopher.

Electron Diffraction Methods of Studying Organic Films*

BY L. H. GERMER

Bell Telephone Laboratories, Inc.
New York, N. Y.

NO attempt has yet been made, so far as I know, to use electron diffraction as a research tool in biology. The newness of electron diffraction, as well as newness of the entire science of biophysics, leads one to suspect that applications to biology may have been simply overlooked. I am not qualified to suggest how such applications can be made. I propose merely to illustrate here the type of information obtainable, by giving briefly the results of studies of some organic substances. It must naturally be left to biologists to say whether or not the technique of these experiments can be applied to much more complex biological materials. If the application can be made it will then, perhaps, be the problem of the physicist to find whether or not the results can be interpreted.

The special field of usefulness of electron diffraction is the investigation of the chemical and crystalline constitutions of very thin films and of surface layers. Biologists are sometimes concerned with the structures of very thin films, and there seems nothing in the nature of biological films which will exclude all of them from investigation by this method. In particular, excessive thinness, which would make a biological film unavailable for examination by x-rays, can place no limitation whatever on its examination by electron diffraction. Films have been studied which are one molecule thick, and from such studies we have deduced the arrangement of molecules in the film and of most of the atoms within the molecules. In another case electron diffraction patterns have revealed long chain molecules folded back and forth upon themselves like yard long railroad tickets; the thickness of the film made up of these molecules is only a small fraction of the length of one molecule. One naturally assumes that information of this type regarding biological films would be of interest and importance.

Experimental Technique

The method of experimentation is quite simple. A narrow beam of electrons is generated within a large vacuum tube. This beam strikes the film to be studied, and electrons scattered from the beam reach a photographic plate which is mounted within the exhausted tube. The scattered electrons leave a record upon the photographic plate; all of the deductions regarding atomic and molecular structure of the film are obtained by studying this record, which is appropriately called an electron diffraction pattern. The laws of electron scattering are similar in many ways to the laws of x-ray scattering, and diffraction patterns produced by electrons resemble those produced by x-rays. We determine arrangements of atoms within crystals, and positions of atoms within molecules, by methods which have become increasingly well known since the first x-ray diffraction experiments in 1912.

Two different ways of mounting specimens to be studied are commonly used. In one of these, called the reflection method, the electron beam strikes at small glancing angle upon the surface of a metal block on which lies the film in which

* Presented at the Symposium on Biophysics, Philadelphia, Pennsylvania, November 4, 5 and 6, 1937.

we are interested. The alternative arrangement is described as the transmission method; in this the film is mounted across a small pinhole, or a narrow slit, and the beam of electrons passes directly through it. A photograph of the standard mounting, which we use in this type of experiment, appears as Fig. 1. In the other figures of this paper are exhibited examples of diffraction patterns obtained from specimens mounted in the two ways.

The Transmission Method

The pattern which appears as Fig. 2 was obtained by passing a beam of electrons normally through a film of polyethylene-sebacate mounted across the slit represented in Fig. 1. Polyethylene-sebacate is a highly polymerized compound, the mean molecular weight of the specimen from which this film was made being about 18,000. The sharp rings of Fig. 2 prove that the material of the film is crystalline, and that the orientations of the crystals are entirely random about the direction of the normal to the film. That, however, the arrangement of crystals is not random about other directions can be readily proved by taking a photograph with the film rotated so that it no longer stands perpendicular to the electron beam; the resulting diffraction pattern is then greatly altered, being made up of arcs instead of continuous rings. This pattern is not reproduced here.

An entirely different alteration of the ring pattern of polyethylene-sebacate is produced by stretching the film before mounting across the slit. In Fig. 3 is shown the pattern obtained from a stretched film, with the stretching direction parallel to the horizontal line through the center of the figure and the electron beam normal to the film. From this pattern and that of Fig. 2 a number of important deductions can be made. Among these are the length of the repeating unit in each of the long molecules, and the orientation of the axes of the molecules in the stretched and in the unstretched films. It is found that in stretched films the molecular axes are parallel to the stretching direction, whereas in unstretched films the axes lie approximately in the plane of the film but have random directions in this plane.

Diffraction patterns produced by films of a different material, gutta-percha, are exhibited in Figs. 4 and 5. Gutta-percha, like polyethylene-sebacate, is a highly polymerized long chain compound, but of a different type. It is unsaturated and contains only carbon and hydrogen, whereas polyethylene-sebacate is saturated and contains oxygen as well as carbon and hydrogen.

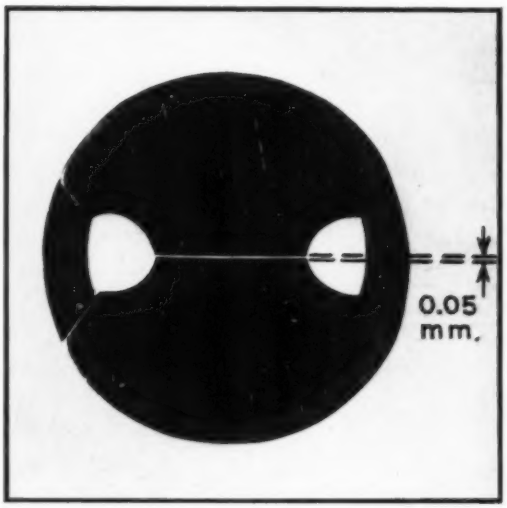


Fig. 1. Photograph of slit used as support for films in experiments by the transmission method.

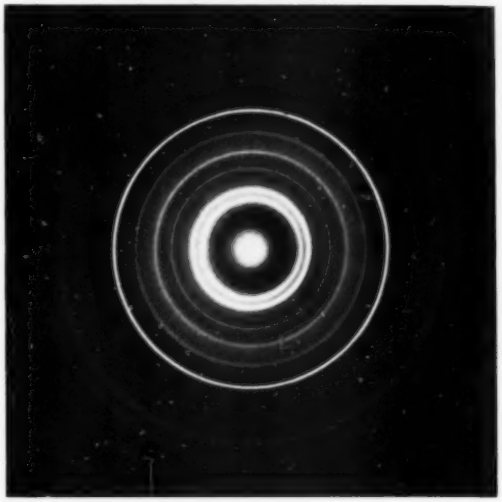


Fig. 2. Electron diffraction pattern from an unstretched film of polyethylene-sebacate, by the transmission method. Electron wave-length, 0.070A.

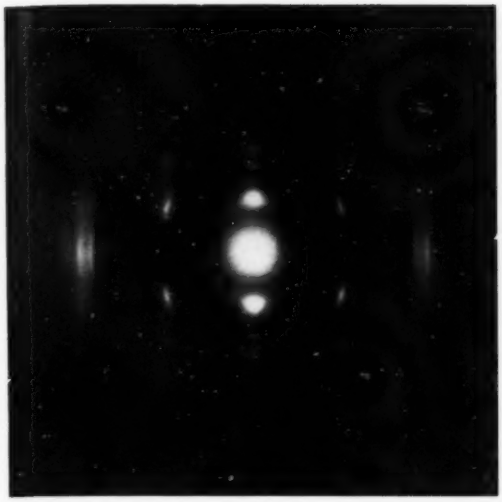


FIG. 3. Diffraction pattern from stretched film of polyethylene-sebacate.

Although Fig. 4 is a pattern of rings like Fig. 2, the interpretation is widely different; the pattern of Fig. 4 is due to long chain molecules the axes of which lie normal to the film surface, rather than parallel to it. Yet in spite of this fact the mean molecular weight of gutta-percha is so great that the average molecule must have a length many times greater than the film thickness. This suggests at once that the molecules are folded back and forth upon themselves normal to the film. The difference between the orientation in thin films of polyethylene-sebacate molecules and gutta-percha molecules is surprising. The reason for the difference is not yet understood.

Although unstretched gutta-percha and polyethylene-sebacate films have entirely different molecular orientations, stretched films have the same. In both materials stretching re-orients the molecules with their axes parallel to the stretching direction. This type of orientation is deduced from Figs. 3 and 5.[1]

The diffraction patterns appearing in Figs. 2–5 were produced by films which lay across the jaws of the slit shown in Fig. 1. Each of the films was somewhat less than 10^{-5} cm in thickness. Films as thick as this are supported very nicely by the jaws of the slit. Sometimes, however, one wishes to examine very much thinner films by the transmission method, or to form films directly upon a supporting foil. As supports for films of excessive thinness we have used foils made from a material known as "Rezoglas."[2] These are much stronger than foils of cellulose nitrate of equal thickness. Rezoglas foils less than 10^{-6} cm thick are easily made, and are self-supporting across the 0.05 mm width of our slit. Such a foil scatters electrons so very weakly that it constitutes a remarkably satisfactory support for other materials. For example, scattering from the Rezoglas support does not interfere appreciably with the diffraction pattern from two molecular layers of barium stearate deposited on the foil, or even with the pattern from two layers of stearic acid. A pattern from such a composite film is shown in Fig. 6; this was produced by two layers of barium stearate molecules deposited on a Rezoglas foil by the Langmuir-Blodgett technique.[3] The diffraction pattern from the underlying Rezoglas cannot be perceived in this reproduction.

THE REFLECTION METHOD

Turning now to the reflection method I shall illustrate the type of results obtainable by referring to a series of experiments on films composed of one or more molecular layers of stearic acid and of barium stearate. In each case the film is prepared, first by cleaning the surface of a chromium plated metal block to remove from it the last layer of greasy organic compounds, and then by dipping the block through a compressed unimolecular layer of stearic acid or of barium stearate floating on water. In this way one single

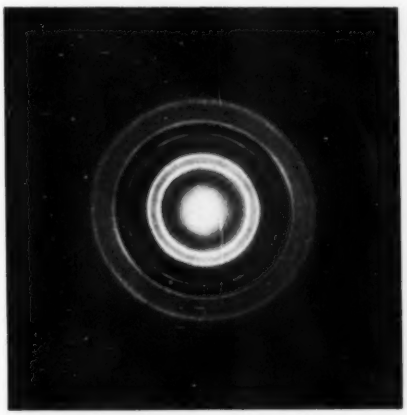


FIG. 4. Pattern from unstretched film of gutta-percha.

molecular layer, or three layers, or any odd number of layers can be built up on the surface. The technique is that developed by Dr. Blodgett.[3]

The pattern shown in Fig. 7 was produced by a single molecular layer of barium stearate, with the primary electron beam grazing the metal surface upon which the layer rested. The diffuse vertical bands constitute the only features of this figure. These bands can be readily interpreted; they can be thought of as the diffraction pattern of single isolated molecules. We are able to conclude that the carbon atoms of each molecule have a planar zigzag arrangement, and that the axes of the molecules are normal to the bands and thus *normal to the surface upon which the molecules rest*. There are almost no regularities in the spacings between neighboring molecules.

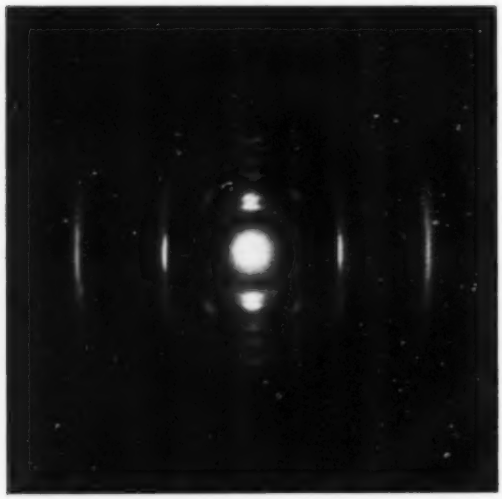


FIG. 5. Pattern from stretched gutta-percha.

In a film of three molecular layers also, barium stearate molecules arrange themselves with their axes normal to the supporting surface, but in this thicker film other regularities occur too. The rather sharp horizontal line segments of Fig. 8 arise from such regularities, and give information regarding the structure of the crystals which are being formed.

Figure 9 is reproduced for comparison with Figs. 7 and 8. This is a diffraction pattern produced by 15 molecular layers of stearic acid.[4] From the sloping bands of Fig. 9 one concludes that stearic acid molecules stand with their

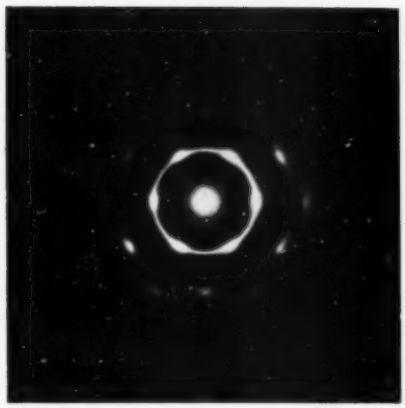


FIG. 6. Diffraction pattern from 2 molecular layers of barium stearate supported by Rezoglas. Thickness of the barium stearate, 5×10^{-7} cm.

axes *inclined* to the plane of the supporting surface. Patterns from smaller numbers of molecular layers of stearic acid also are characterized by sloping bands, except for the unique case of a single layer only. The bands which constitute the pattern of a single molecular layer of stearic acid are not inclined, but are parallel to the surface as in the patterns of barium stearate. The stearic acid molecules of a single layer stand with their axes normal to the surface, although subsequent layers stand inclined. The conclusion is almost unavoidable that the first layer of stearic acid combines at once with a metal surface to form a stearate.

FIG. 7. Diffraction pattern from a single layer of barium stearate molecules, by the reflection method.

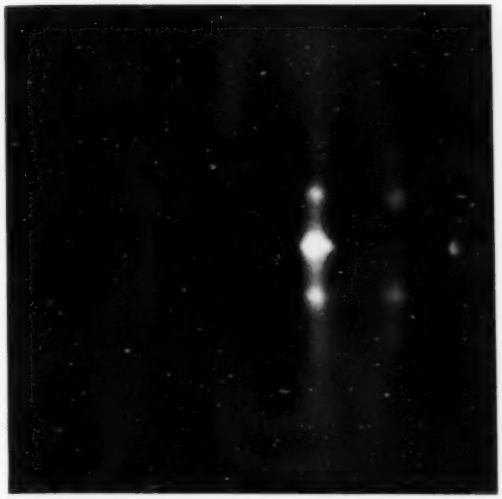


FIG. 8. Pattern from 3 molecular layers of barium stearate.

Upon this original layer other layers can then exist as uncombined acid.

These experiments seem to me to be significant in two ways; in the first place, they reveal the arrangement and structure of organic molecules upon surfaces, in the second place, we believe that they reveal chemical combination between a surface and the first layer of acid molecules. These results have interesting bearing upon theories of lubrication, in particular the nature of "oiliness" and boundary lubrication. Boundary lubrication deals with surface phenomena, as do the preponderance of biological processes.

Conclusion

Applications to biology of the experimental methods illustrated here are by no means obvious. Any application can be divided into two steps; in the first place, the material to be studied must be prepared in a form suitable for examination; in the second place, the diffraction pattern obtained must be interpretable. In view of the fact that studies of biological materials may possibly be made, it seems that some of the difficulties should be pointed out. Under the first step of the procedure, preparation of the specimen, a major difficulty may arise from the fact that the specimen must be stable in vacuum. It has been suggested that some materials can be frozen quickly in liquid air, and introduced into the vacuum chamber in the frozen condition. Even if this can be carried out without change in the nature of the substance, experimental difficulties are serious. And difficulties which are no less serious will arise in the second step of the procedure, the interpretation of the diffraction pattern. In the relatively easy experiments reported here, successful interpretation has been favored by comparative simplicity of the materials studied; all of them have been long chain compounds which yield diffraction patterns of a simple type.

I believe that some materials of a very complex nature will ultimately be successfully studied by these methods. I do not wish, however, to leave the impression that the work will be unduly easy.

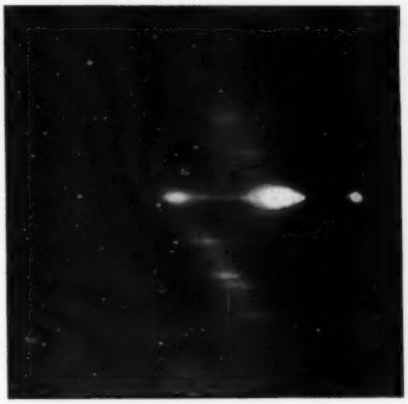


FIG. 9. Pattern from 15 molecular layers of stearic acid.

References

1. It is expected that a complete account of the investigation of polyethylene-sebacate, gutta-percha, and some other highly polymerized substances will soon be published by Mr. K. H. Storks of these laboratories.
2. This material was obtained from Advance Solvents and Chemical Corporation, 245 Fifth Avenue, New York City.
3. Blodgett, J. Am. Chem. Soc. **57**, 1007–1022 (1935).
4. Figure 9 appears as Fig. 1 in a paper by Germer and Storks, Proc. Nat. Acad. **23**, 390–397 (1937).

The Biophysics and Biochemistry of Viruses*

BY W. M. STANLEY

The Rockefeller Institute for Medical Research
Princeton, New Jersey

WHEN the invitation to participate in this symposium was tendered, it was suggested that I speak about the nature and organization of elementary biological systems. It became necessary, therefore, to select some biological system as the subject for this discussion. As we go from the more complex organisms towards the more elementary, we finally arrive at the bacteria as the simplest type which is universally accepted as truly representative of living things. However, there would be little merit in selecting the bacteria as the subject of the present discussion, for their reproduction, metabolism, organization, structure, mutation, and so forth have been rather extensively studied and discussed, and it is generally accepted that in most respects the bacteria are similar to the more complex biological systems. We are forced, therefore, to look still lower in our scale for a suitable subject. Just below the bacteria with respect to size are the viruses, mysterious purveyors of disease, about which we hear so much today. If we pass over the viruses and go still lower we come to the enzymes. For many years enzymatic activity was considered to be a vital activity which could take place only within living cells. This idea eventually had to be revised, for first one and then several of the enzymatic reactions were found to take place *in vitro* when suitable conditions were achieved. This was followed by the isolation of several of the active agents in the form of crystalline proteins.[1–7] It is now becoming generally accepted that the enzyme proteins, such as pepsin, trypsin, and urease, are in fact the enzymes themselves, and that the enzymatic activity is a specific property of the protein molecules. It is interesting to note that these protein molecules have been found to undergo certain chemical reactions[8] in much the same manner that ordinary molecules, such as, let us say, those of the amino acids, undergo chemical reactions. Despite the tremendous importance of enzymatic activity in practically all types of biological systems, it is of itself not a vital activity, but is in reality merely the manifestation of the chemical properties of certain molecules. Enzymatic reactions are chemical reactions and appear to be governed by the ordinary laws of chemistry. The enzymes are, therefore, ordinary molecules and according to present standards should not be regarded as true representatives of a very simple type of biological system.

We are forced, therefore, to retrace our steps and to consider the possibility of viruses as elementary biological systems. A discussion of viruses is peculiarly appropriate in this symposium, because the discovery of their very existence was accomplished by physical means, and more especially because, as I shall point out later, physical methods have become of great importance in isolating and studying the viruses. The existence of this great group of disease-producing agents that we now know as viruses

* Presented at a Symposium on Biophysics held in Philadelphia, Pennsylvania, on November 4, 5 and 6, 1937.

was first demonstrated in 1892 by Iwanowski.[9] He took the sap from a plant affected with the tobacco mosaic disease, passed it through a filter which held back all of the living organisms then known, and found that the filtrate was infectious and would cause the tobacco mosaic disease when injected into a normal plant. Later Beijerinck[10] ruled out the possibility of a toxin by showing that the disease could be carried repeatedly from plant to plant by means of such filtrates. This simple physical separation permitted the recognition of tobacco mosaic virus as an infectious disease-producing entity differing at least in size from the known living organisms. Since that time similar experiments have enabled us to recognize a whole host of viruses affecting not only plants but bacteria, animals, and man as well. Diseases such as the several mosaic and yellows diseases of plants, rabies, dog distemper, foot-and-mouth disease, equine encephalitis, poliomyelitis, yellow fever, the St. Louis type of encephalitis, certain types of tumorous growths in animals, and even the common cold are caused by viruses.

Now let us consider some of the characteristics of these infectious disease-producing agents that are still smaller than bacteria. Perhaps their most striking property is growth, or reproduction, or multiplication. For example, we may take one unit of a given virus and infect a suitable normal host, and within a short time millions of units of virus may be demonstrated in the formerly virus-free host. Under such conditions the virus grows or reproduces. However, if the virus is held in the test tube or injected into nonsusceptible hosts, nothing happens, disease is not caused, the amount of virus does not increase. Viruses have been found to reproduce, therefore, only within the living cells of certain susceptible hosts. Still another property of viruses is their ability to change and adapt themselves to new conditions and surroundings and thus to mutate and to give rise to new strains. By proper manipulation you can change one virus strain into a new and different strain. For example, smallpox virus passed from man to monkeys to calves becomes vaccine virus, which, when passed back to man, causes not smallpox but a mild disease, vaccinia. You are all familiar with vaccine virus, for it is used to vaccinate or protect you against smallpox. Because viruses may reproduce under certain conditions, because they are specific in that certain viruses occur only in certain hosts, and because they may change or mutate, most workers in the virus field have chosen to regard viruses as small living organisms, for the properties just mentioned are characteristic of living things. However, there have been certain objections to the complete acceptance of this idea. The fact that it has proved impossible to cultivate viruses in the absence of living cells has been frequently cited as evidence that viruses differ in nature from bacteria. However, the property of obligate parasitism is in fact not peculiar to viruses, for many organisms are known to be obligate parasites and to require the presence of certain living cells for reproduction. A more pertinent fact is that certain viruses are smaller than some known molecules. For example, yellow fever virus is considered to be about 22 mμ in diameter[11, 12] and the virus causing foot-and-mouth disease only 10 mμ in diameter,[13] whereas the diameter of certain of the hemocyanin protein molecules is 24 mμ.[14] The difficulties in considering these small viruses as living organisms is obvious, for we have always assumed that a certain degree of complexity was necessary for life. It is problematical whether such small entities no larger than ordinary protein molecules can support that degree of complexity. Because of such objections, which were early recognized, bacteriologists and pathologists voluntarily separated these minute disease-producing entities from ordinary bacteria, placed them in a separate group, and called them viruses. Since the viruses as a group are considerably smaller than bacteria, yet possess many properties considered to be characteristic of living things, they qualify eminently as a representative of an elementary biological system. If further qualification were necessary, the apparent parallelism between the properties and characteristics of viruses and of genes could be cited. Viruses were, therefore, selected as the subject of this discussion.

Now let us consider what has been revealed about the nature and organization of viruses by biochemical and biophysical studies. We may well divide these studies and consider first only those which took place before 1935. Before that

time viruses were regarded variously as very small ordinary bacteria, as some new type of invisible living organism, as toxins, as enzymes, as non-nitrogen-containing and as nitrogen-containing chemical principles.[15] They had not been isolated as such and were known only by the biological effects they produced. They were first recognized and continued to be recognized only by the diseases which they caused. This state of affairs seriously handicapped investigators, for the agents with which they were working were apparently even more elusive than the molecules with which the physicist and chemist work. The latter are usually able to isolate the molecules of the substance they work with so that they can handle, weigh, and ascribe physical and chemical properties to it. All of this was lacking with the viruses, for even their general nature was unknown. In an attempt to gain information concerning the nature of viruses, many studies on the effect of various chemical agents on virus activity were made, but, due to the failure to differentiate between effect of chemical on virus and effect of chemical on the host used to test for virus, the results were frequently contradictory and only served to confuse matters because of the erroneous impressions which resulted. The physical studies on viruses centered around the measurement of size by means of filtration through graded collodion membranes.[16] With the exception of measurements on certain viruses existing in low concentration, the filtration results were quite good and served to give an excellent idea of the relative size of many of the viruses. However, because a tangible entity was not available, most of the physical measurements were dependent upon the biological measurements of virus activity. This was a severe handicap, for the error of the biological measurements was about a thousand percent.

A turn of events occurred early in 1935 when it was found possible by chemical means to isolate tobacco mosaic virus in the form of a crystallizable high molecular weight protein.[17] The first impression of the unusually high molecular weight of the protein was obtained by diffusion experiments. However, the diffusion method is not especially well suited for such huge molecules, so a sample was sent to Professor Svedberg for an ultracentrifugal analysis. He found the protein to have a sedimentation constant of about 200 and a probable molecular weight of about 17 million.[18] I should like to emphasize the great importance of the analytical ultracentrifuge in work with the virus proteins. The ultracentrifuge was first designed by Svedberg for the specific purpose of measuring the rates of sedimentation of large colloidal particles, but he has perfected the equipment so that measurements on proteins may be made with great accuracy. In the case of the virus proteins, the ultracentrifugal method is superior to diffusion or osmotic pressure, and we would have been severely handicapped had not the equipment been made available to us first by Dr. Svedberg and later by Dr. Wyckoff.

A sufficient quantity of tobacco mosaic virus protein was isolated to permit thorough chemical and physical studies of its properties. I will take up the results of some of these studies later, and at this point I will simply say that $2\frac{1}{2}$ years' study has demonstrated only that the protein is homogeneous from the physical, chemical, and biological standpoints, and that there is no reason for assuming other than that this protein is tobacco mosaic virus (see reference 19 for reviews and bibliographies). One of the important considerations that came up almost immediately upon the isolation of this heavy protein was the question as to whether or not tobacco mosaic virus was a special case and differed from all other viruses. It is interesting to note that previously there had been no reason to differentiate tobacco mosaic from the other viruses, but that with the isolation of the heavy protein the question was raised and became of paramount importance to many people. It became necessary, therefore, to determine whether or not similar heavy proteins existed in the cases of other virus diseases. Plants diseased with strains of tobacco mosaic virus were first examined by means of the chemical method used for the isolation of tobacco mosaic virus protein, and in every instance specific heavy proteins were isolated. These proteins were different and were characteristic of the diseases which they produced. However, when these same methods were applied to plants diseased with some of the less stable viruses, such as latent mosaic of potato, ring spot, severe etch,

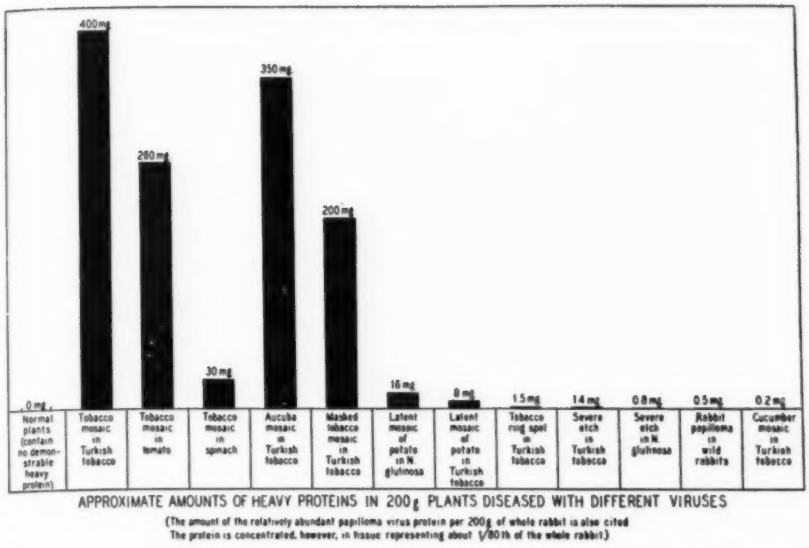

400mg
260mg
30mg
350mg
200mg
16mg
8mg
1.5mg
1.4mg
0.8mg
0.5mg
0.2mg
0 mg.
Normal plants (contain no demonstrable heavy protein)
Tobacco mosaic in Turkish tobacco
Tobacco mosaic in tomato
Tobacco mosaic in spinach
Aucuba mosaic in Turkish tobacco
Masked tobacco mosaic in Turkish tobacco
Latent mosaic of potato in N. glutinosa
Latent mosaic of potato in Turkish tobacco
Tobacco ring spot in Turkish tobacco
Severe etch in Turkish tobacco
Severe etch in N. glutinosa
Rabbit papilloma in wild rabbits
Cucumber mosaic in Turkish tobacco
APPROXIMATE AMOUNTS OF HEAVY PROTEINS IN 200 g PLANTS DISEASED WITH DIFFERENT VIRUSES
(The amount of the relatively abundant papilloma virus protein per 200 g. of whole rabbit is also cited
The protein is concentrated, however, in tissue representing about 1/80th of the whole rabbit)

FIG. 1. Approximate amounts of heavy proteins in 200 g of plants diseased with different viruses.

and cucumber mosaic viruses, it was found impossible to isolate heavy proteins by this chemical method. These viruses are considerably less stable and are much less infectious than tobacco mosaic virus. Their dilution end points are, in some instances, only one ten thousandth that of tobacco mosaic virus. It seemed possible, therefore, that the concentration of these viruses in the host might be thousands of times less than that of tobacco mosaic virus and hence, although there might be heavy proteins present, that the chemical method might not be sufficiently specific to separate such heavy proteins from the great bulk of low molecular weight proteins.

Happily, about this time the air-driven quantity ultracentrifuge was developed.[20–24] One of these centrifuges holding 130 cc and capable of being run at a speed which sediments over 99.9 percent of tobacco mosaic virus protein from solution within 90 minutes was made available by Dr. Wyckoff. Since practically all of the proteins and other substances that we know about have relatively low molecular weights, it was obvious that, if viruses were heavy proteins, it should be possible to separate them from the low molecular weight material by differential ultracentrifugation. As a matter of fact, centrifugation seemed to be the one specific method for the separation, purification, and isolation of heavy proteins. Extracts of Turkish tobacco plants diseased, respectively, with latent mosaic of potato, tobacco ring spot, severe etch, and cucumber mosaic viruses were subjected to differential ultracentrifugation, and in every instance it was possible to demonstrate the presence of a heavy protein. Dr. Wyckoff examined in the analytical ultracentrifuge the heavy proteins which were isolated by this method and found the latent mosaic virus protein to have a sedimentation constant of about 113, tobacco ring spot virus protein about 110, severe etch virus protein about 180, and cucumber mosaic virus protein about 120. These heavy proteins were found to possess the properties of the respective viruses and to have different physical and chemical properties. The purely physical method of differential centrifugation thus succeeded where chemical methods had failed, and enabled the isolation of a whole series of heavy virus proteins.

One important finding resulting from this work was the demonstration of the vast difference in the concentration of the different virus proteins in the same host. Fig. 1 shows the relative amounts of heavy proteins in 200 g tobacco plants diseased with different viruses. It may be seen that normal plants contain no demonstrable heavy protein, that tobacco mosaic diseased Turkish tobacco plants contain 400 mg of virus protein per 200 g of plant, or about one part of virus protein per 500 parts of plant. However, diseased tomato plants were found by Dr. Loring to contain only 260 mg and diseased spinach plants only 30 mg. The fact that the concentration of tobacco mosaic virus protein varies with the host plant is interesting and perhaps significant. It indicates that in different plants the level which the virus protein reaches may be dependent upon a mechanism characteristic of that plant. It may also be seen that the amount of virus protein in a given host varies with the

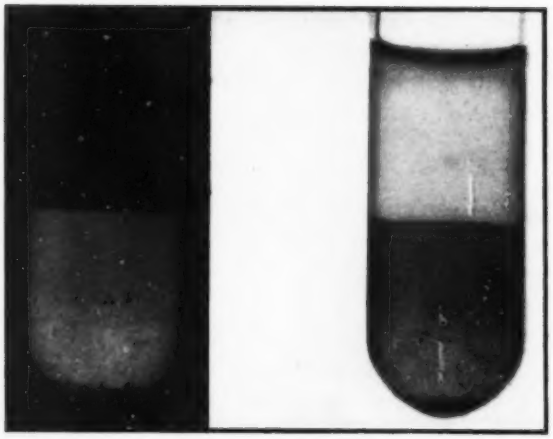
strain of the virus; thus in Turkish tobacco aucuba mosaic virus protein reaches a level of only 350 mg instead of 400, and the masked strain a level of only 200 mg. Now if you will contrast the amounts of tobacco mosaic virus protein with the amounts of other virus proteins, you will see at once why the chemical method of isolation was so successful in the case of tobacco mosaic virus and its strains and not successful in the cases of the other viruses. The amount of tobacco mosaic virus protein is very large when compared to the amounts of the other virus proteins. Actually one wonders how the presence of such vast amounts of this unusual heavy protein in diseased Turkish tobacco plants remained undiscovered for so many years. It is extremely fortunate, however, that nature provided us with one virus in such amounts that we could not fail to discover it, for it was by means of tobacco mosaic virus protein that we entered upon the whole field of virus proteins. The lessons learned from tobacco mosaic virus protein have been very valuable to us in our work with other viruses. Had we been forced to do the exploratory work with, let us say, cucumber mosaic virus, it is probable that many years would have elapsed before the existence of the heavy virus proteins would have been discovered. Now we see that through tobacco mosaic virus protein and the quantity ultracentrifuge we have been able to discover a whole host of heavy virus proteins, the very existence of which was not only unknown but even unpredicted 3 years ago. Although before the isolation of the plant virus proteins there was no objection to regarding the viruses affecting plants and the viruses affecting animals as being quite similar, immediately following the isolation of tobacco mosaic virus protein a definite tendency to regard the plant viruses as being quite different from animal viruses became apparent. It became desirable, therefore, to determine whether or not heavy proteins existed in the case of animal virus diseases.

There was under investigation at this time in the Princeton laboratories of The Rockefeller Institute a highly infectious relatively stable virus which Dr. Shope[25] had discovered to cause warts or papillomas in rabbits. Dr. Beard subjected an extract of the wart tissue from rabbits to the centrifugation technique and was able to isolate a protein which Dr. Wyckoff found to be homogeneous and to have a sedimentation constant of about 250 and a probable molecular weight of over 20 million.[26] This protein caused papillomas when injected into rabbits and was found to be several thousand times more active than the tissue from which it was obtained, only 10^{-8} g being required to cause infection. Insofar as determined, this protein possesses the properties of the virus and there is at this time no reason to assume that it is other than the virus. Unusual interest surrounds this virus protein, for Rous and Beard[27] have found that the papillomas caused by it usually progress and become cancers. At this same time work was being conducted with the virus of equine encephalitis, a virus which differs markedly from the papilloma virus in that it affects the central nervous system and is very unstable. This virus is of interest because it resembles in many respects poliomyelitis virus. Extracts of tissues diseased with equine encephalitis virus were subjected to the centrifugation technique and there was obtained a material which Dr. Wyckoff found to have a sedimentation constant of about 245 and a probable molecular weight of about 25 million.[28] This virus is quite unstable and the heavy material was found to disintegrate rapidly into low molecular weight material on standing at room temperature. The success of these first two investigations in the very difficult field of animal viruses is a tribute to the effectiveness of the centrifugation technique. The mere centrifugation of a solution containing virus first at a speed sufficient to throw down large particles and colloidal material, and then in new tubes at a speed sufficient to sediment the virus and thus to separate the virus first from colloidal matter and then from low molecular weight material may appear quite simple and elementary. However, it is this simple physical technique which has enabled the isolation of unstable virus proteins and of virus proteins existing in low concentration and which has permitted us to capitalize to the fullest extent on the knowledge first gained by the isolation of tobacco mosaic virus protein by chemical means. Thus, the physical method of ultrafiltration first revealed the existence of this

great group of disease-producing entities, and now through knowledge first gained by chemical isolation the physical method of centrifugation has enabled the actual isolation of many of these entities.

Now, in order to learn something about the nature and organization of these infectious high molecular weight proteins, it became necessary to study their physical and chemical properties. It is obvious that such studies could be carried out only with great difficulty on unstable virus proteins or on those existing in low concentration. For example, we have obtained cucumber mosaic virus protein at best in one milligram lots, and it was so unstable that within a very few hours it had disintegrated. You can imagine the difficulties involved in making almost any kind of a determination with it. However, because of the knowledge of the existence of these other plant and animal virus proteins, we were able to return to a study of the physical and chemical properties of the abundant and relatively stable tobacco mosaic virus protein, confident that the knowledge gained would be typical and useful in the study of the less abundant and less stable virus proteins. Now permit me to review briefly what we have found out about the physical and chemical properties of tobacco mosaic virus protein.[19, 29] The protein obtained from many different batches of starting material had the same chemical composition, isoelectric point, optical rotation, biological activity, sedimentation constant, crystalline x-ray diffraction pattern, electrophoretic behavior, and immunological properties. As it exists in the plant cells, the protein is homogeneous with respect to size, and if handled carefully can be isolated in a homogeneous condition. It tends to become inhomogeneous when treated with salt or when subjected to changes of hydrogen ion concentration. Protein homogeneous or inhomogeneous with respect to size was found to be homogeneous with respect to electrophoretic mobility. It was found impossible to separate virus activity from protein by means of filtration through collodion or other types of filters, by drastic fractional crystallization, or by centrifugation of the high molecular weight protein from solution under a variety of conditions. The absorption spectrum of the protein was found to agree essentially with the destruction spectrum of the virus activity. Denaturation of the protein by several different methods was always accompanied by loss of virus activity. Treatment with hydrogen peroxide, formaldehyde, nitrous acid, or ultraviolet light was found to produce inactive native proteins that, although slightly altered, still retained certain chemical and serological properties characteristic of the virus protein. It was found possible to separate mixtures of such inactive protein and active protein by mere fractional crystallization. Under specified conditions the protein possessed a definite and characteristic solubility. The density of the protein was found to be 1.33, much higher than the value of 1.1 which is usually given for bacteria.

One of the properties of the protein of interest to physicists is the marked double refraction of flow which is exhibited by its solutions.[30] This phenomenon is even more marked than that shown by the classic vanadium pentoxide sols. It is an indication, of course, that the molecules of the protein are rod-shaped and tend to orientate when made to flow. If the dissymmetry of the molecules is of an extreme nature such as that ascribed to the molecules of myosin, it would be expected that the solutions of tobacco mosaic virus protein would be unusually viscous. However, the viscosity was found to be much less than that of myosin and to lie between that of gelatin and egg albumin. It seems likely that the molecules of tobacco mosaic virus protein are not thread-like but are, let us say, of the order of about 10 times as long as their cross section. This dissymmetry is used as the basis for explaining the liquid crystalline state of the protein which has been studied by Bawden and Pirie.[31] If a rather concentrated solution of tobacco mosaic virus protein is allowed to stand for some time, it will separate into 2 layers which have different solids contents and different appearances. The upper layer is the more dilute and shows double refraction only when made to flow. The lower layer is the more concentrated and is spontaneously birefringent. The two layers may be mixed and after standing will again separate. As may be seen from Fig. 2, which is a reproduction of a photograph, in polarized light, of a solution of tobacco mosaic virus protein that has been allowed to stand, the line of demarcation between



A B

FIG. 2. Photograph in polarized light of a solution of tobacco mosaic virus protein that has been allowed to stand. (*A*) Test tube with crossed Polaroid plates on opposite sides. (*B*) Test tube with parallel Polaroid plates on opposite sides. The lower layer is spontaneously birefringent, whereas the upper layer is not. (Photograph by J. A. Carlile.)

the two layers may be very sharp. There is no reason to assume that the layering is due to the presence of two different components, for lower layer material may be obtained by concentrating upper layer material and upper layer material may be obtained by diluting lower layer material. Although Bawden and Pirie found no difference between the virus activity of the upper and lower layers, they consider the lower layer to be more pure. The layering phenomenon is quite interesting and apparently results from the force of gravity. Bawden and Pirie assume that when a certain concentration is reached the virus particles are so close together that rotation about the two shorter axes becomes impossible, although translational motion is unimpeded, and that the fluid then consists of a three-dimensional mosaic of regions arranged at random to each other, but in each of which all of the rod-shaped particles will lie approximately parallel. Bawden and co-workers[32, 33] consider that the orientation of the molecules in the crystals which we have described is similar to that of the molecules in the lower layer. They regard our crystals not as true crystals, but as paracrystals having only a two-dimensional regularity. They state that this degree of regularity has not been described previously but that it has been predicted theoretically as one of the types of liquid crystal. I am not qualified to discuss this question, but I can say that, insofar as the manipulation in the laboratory is concerned, our crystals behave as crystals. They can be dissolved and recrystallized and mixtures can be separated by fractional crystallization. However, I am quite willing to leave to the crystallographer and the x-ray specialist the question as to whether or not these crystals are true or paracrystals and the interesting claim of the English workers that the crystalline virus protein represents a new type of crystalline form.

Another of the properties of especial interest is the ability of the protein to change or become altered or, shall we say, mutate and give rise not to the tobacco mosaic disease but to a new disease. We know that when this change takes place there is an actual change in the protein, for the protein isolated in the case of the new disease is not tobacco mosaic virus protein but a new and different although related protein. Thus, aucuba mosaic is a strain of tobacco mosaic, yet aucuba mosaic virus protein can be distinguished from tobacco mosaic virus protein by means of its physical and chemical properties. If this change or mutation occurred only within living cells, it would be as mysterious and as difficult to attack as the mutation of genes, but preliminary experiments indicate that a change can be caused by suitable treatment of the protein while it is in a test tube. This brings it into the realm of chemistry and we hope therefore to be able eventually to correlate such changes with certain definite changes in the structure of the protein molecule.

Now if you have not noticed it before, you most certainly noticed just now, when I spoke of mutation and molecules almost in the same breath, that I have been referring to the virus proteins as molecules and not as organisms. I have done this because I have been discussing their physical and chemical properties, and the physical and chemical properties of the virus proteins are the physical and chemical properties of molecules.[29] There is no reason to segregate the virus proteins because of their size, for despite their tremendous size they have the ordinary chemical and physical properties of molecules and these properties are a direct consequence of

their chemical structure. An objection to the final and complete acceptance of virus proteins as molecules is that, although an analogy of a protein reproducing itself *in vitro* is known,[34] the autocatalytic production of a protein *in vivo* is today not an accepted chemical reaction. Immediately we are forced inside of a living cell, we are faced with a set of conditions that we have not fathomed as yet. The virus protein is apparently imbued with something of which today we are ignorant. Its biological activity as evidenced by the ability of reproduction when within the cells of certain hosts is a property that we hitherto have not ascribed to ordinary molecules. It is a property of an elementary biological system, and, since it is possessed by something which the chemist regards as a molecule, the virus proteins appear to be the elementary biological system. Today we are face to face with this problem. Rather than arbitrarily conclude that the virus proteins are either molecules or organisms, I feel that we should be less interested in their classification and more interested in their biological activity, for this property of reproduction, of infectiousness if you please, is the distinguishing and characterizing property of the virus proteins. Knowledge of the mechanism by means of which a given huge molecule-like protein, when introduced into a certain cell, is able to bring about the organization of the amino acids and cellular constituents with the formation of exact duplicates of the introduced entity is of tremendous importance, for this mechanism is the basis of biological activity. The virus proteins, because of their abundance and stability, furnish us with adequate working material for a thorough study of this mechanism. It is possible that the solution may lie outside of the chemical field and in the realm of physics, and it is to this end that I hope for a boom of interest in the physics of macromolecules comparable to that recently enjoyed by nuclear physics. The ultimate solution of the general virus problem will probably come about through the combined efforts of workers in several fields.

Bibliography

1. J. B. Sumner, J. Biol. Chem. **69**, 435 (1926).
2. J. H. Northrop, J. Gen. Physiol. **13**, 739 (1930).
3. M. L. Caldwell, L. E. Booker and H. C. Sherman. Science **74**, 37 (1931).
4. J. H. Northrop, and M. Kunitz, J. Gen. Physiol. **16**, 267 (1932).
5. H. Theorell, Biochem. Z. **272**, 155 (1934).
6. M. Kunitz and J. H. Northrop, J. Gen. Physiol. **18**, 433 (1935).
7. M. L. Anson, Science **81**, 467 (1935).
8. R. M. Herriott, J. Gen. Physiol. **19**, 283 (1935).
9. D. Iwanowski, Bull. Acad. Imp. Sci. St. Petersburg, n.s. III, **35**, 67 (1892).
10. M. W. Beijerinck, Verhandel. Konink. Akad. Wetensch. Amsterdam II, **6**, 1 (1898).
11. G. M. Findlay and J. C. Broom, Brit. J. Exp. Path. **14**, 391 (1933).
12. J. H. Bauer and T. P. Hughes, Am. J. Hyg. **21**, 101 (1935).
13. I. A. Galloway and W. J. Elford, Brit. J. Exp. Path. **12**, 407 (1931).
14. I. B. Eriksson-Quensel and T. Svedberg, Biol. Bull. **71**, 498 (1936).
15. T. M. Rivers, Physiol. Rev. **12**, 423 (1932).
16. W. J. Elford, J. Path. Bact. **34**, 505 (1931).
17. W. M. Stanley, Science **81**, 644 (1935).
18. I. B. Eriksson-Quensel and T. Svedberg, J. Am. Chem. Soc. **58**, 1863 (1936).
19. W. M. Stanley, Am. J. Bot. **24**, 59 (1937); Erg. d. Physiol. **39**, 294 (1937).
20. E. Henriot and E. Huguenard, Comptes rendus Acad. Sci. **180**, 1392 (1925).
21. J. W. Beams and E. G. Pickels, Rev. Sci. Inst. **6**, 299 (1935).
22. J. Biscoe, E. G. Pickels, and R. W. G. Wyckoff, Rev. Sci. Inst. **7**, 246 (1936); J. Exp. Med. **64**, 39 (1936).
23. J. H. Bauer and E. G. Pickels. J. Exp. Med. **64**, 503 (1936); **65**, 565 (1937).
24. R. W. G. Wyckoff and J. B. Lagsdin, Rev. Sci. Inst. **8**, 74 (1937).
25. R. E. Shope, J. Exp. Med. **58**, 607 (1933).
26. J. W. Beard, and R. W. G. Wyckoff, Science **85**, 201 (1937).
27. P. Rous and J. W. Beard, J. Exp. Med. **62**, 523 (1935).
28. R. W. G. Wyckoff, Proc. Soc. Exp. Biol. and Med. **36**, 771 (1937).
29. W. M. Stanley, J. Phys. Chem. **42**, 55 (1938).
30. W. N. Takahashi and T. E. Rawlins, Science **81**, 299 (1935); **85**, 103 (1937).
31. F. C. Bawden and N. W. Pirie, Proc. Roy. Soc. **B123**, 274 (1937).
32. F. C. Bawden, N. W. Pirie, J. D. Bernal, and I. Fankuchen, Nature **138**, 1051 (1936).
33. J. D. Bernal and I. Fankuchen, Nature **139**, 923 (1937).
34. M. Kunitz and J. H. Northrop, Science **80**, 505 (1934).

The Photochemical Basis of Vision*†

BY SELIG HECHT

Laboratory of Biophysics, Columbia University
New York, New York

I. Historical Approach

THE photochemical investigation of vision can be approached from different points of view. In fact, in its development it has been considered from several points of view, and perhaps the best way of summarizing our present knowledge of the subject is to record its historical growth.

We speak easily of the photochemistry of the visual process, and so commonplace does this seem, that it comes as a surprise to learn that the idea is no more than a hundred years old. Almost immediately after the discovery by Niépce and by Daguerre of the fixation of a photochemically produced image on a photographic plate, Moser[30] suggested that the action of light on the retina was similar to the action of light on photosensitive materials generally. Such an idea of retinal action was accepted by many physiologists, though the evidence for it was no more than the analogy with the photographic plate. Since then the photochemical basis of vision has been much studied, and now there exists a body of information about it which is one of the substantial structures of modern physiology.

II. Visual Purple

It would seem that the most obvious approach to the photochemical investigation of vision is to examine the retina for gross color changes in the light and dark, and to make various extracts of it for testing under the same conditions. However, after Moser's suggestion, it required thirty years before anyone performed such experiments. Then in 1876 Boll[5] made the dramatic discovery that the freshly removed frog's retina is purple, and that this color is bleached by light. Even more exciting was Kühne's demonstration the following year that the purple color of the retina is due to a substance, which Kühne called visual purple and which he soon learned to extract from the retina into solution. Boll and Kühne, particularly Kühne, did a number of beautiful researches with visual purple, which raised high hopes for a speedy solution of the problems of vision. They showed, for example, that actual pictures of outside configurations like windows and houses could be made on the retinas of frogs and rabbits, and Kühne found means of fixing these pictures so that they could be demonstrated publicly.

Kühne (1879)[28] studied visual purple both in the retina and in solution. He found that visual purple, as it bleached, always passed through a yellow stage which seemed to be formed from the visual purple, and that this visual yellow was converted by the further action of light into a comparatively colorless substance, visual white. Kühne compared the effectiveness of different parts of the spectrum in bleaching visual purple, and showed that the extreme ends of the spectrum were much less photochemically active than the middle wave-lengths. Of course, Kühne had no means of making energy measurements, and his results were very rough indeed. But they indicated that the effectiveness of the spectrum in bleaching visual purple has about the same distribution as its effectiveness in vision. Kühne was able to show that visual purple when bleached in the retina could regenerate its color by itself and even more rapidly by being replaced

* Presented at the Symposium on Biophysics, Philadelphia, Pennsylvania, November 4, 5 and 6, 1937.

† Publication assisted by the Ernest Kempton Adams Fund for Physical Research of Columbia University.

on the pigment layer which lies behind it in the optic cup. He found that visual purple could be regenerated either directly from its products of decomposition or by the addition of fresh material to the retina. Moreover, he reported that it could be regenerated even in some solutions.

In addition to its sensitivity to light, visual purple is affected by various chemical substances. Materials like ether, alcohol, formaldehyde, and all acids destroy its color. On the other hand Kühne found that ammonia, sodium carbonate, sodium chloride, alum, hydrogen peroxide, urea, and even trypsin digestion left visual purple untouched. Most surprising of all was the clear ineffectiveness of vigorous oxidizing and reducing agents.

Organic and physiological chemistry was not far enough advanced then for Kühne to know what these various destructions and immunities indicated. We know now that many native proteins have their properties changed by the things which destroy visual purple. In this connection the most indicative measurements which Kühne made concern the destruction of the color in the dark at different temperatures. His measurements show an 8-fold increase in the velocity of thermal destruction of visual purple for a rise of 10°C in temperature. This is a high value, and in terms of modern knowledge also points to the protein nature of visual purple.

III. Interregnum

Kühne's work with visual purple, though it was exciting and dramatic, commanded the general interest for a short while only. At the beginning it had raised lively expectations for an understanding of the problems of vision. But the results were disappointing, and in a few years the work lapsed. The causes of this disappointment are worth examining, because they tell us something about the irregular progress of scientific investigation.

In the first place, it was difficult to see any precise relationship between the properties of vision and of visual purple. Except for the very rough spectrum measurements, no similarities existed. The reasons for this are that the properties of vision itself had not been measured with precision, and that no adequate spectrophotometric means existed for the exact evaluation of visual purple.

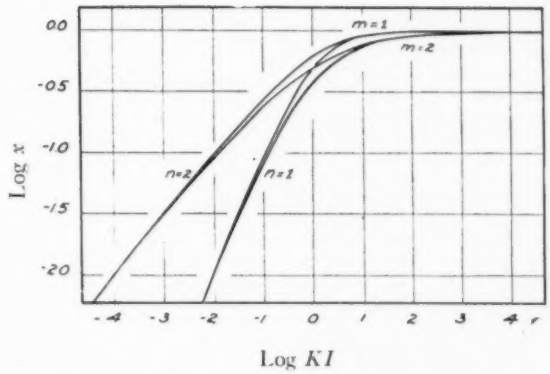


FIG. 1. The steady state Eq. (3) plotted when m and n are each 1 and 2. Because of the log plot the shape of the curves remains constant regardless of the values of K and a.

In the second place, visual purple was found only in the rods and could not be demonstrated in the cones. The difference between these two retinal receptors was not clearly recognized as yet, though it had been pointed out by Schultze in 1866.[31] Nevertheless, the fact that the fovea contained cones alone and that the cones contained no visual purple discouraged one from a belief in the usefulness of this substance as a necessary step in the visual process.

In the third place, Kühne himself demonstrated that it was possible to see, and see well, even when the visual purple was completely bleached in the eye. Frogs placed in bright sunlight showed not a trace of visual purple in the retina; and still, such frogs caught flies with their usual skill, and themselves avoided being caught with their customary agility.

This first and naive approach to the photochemistry of vision therefore apparently yielded little. As a result, the whole subject was almost ignored for many years after Kühne. Koenig (1894)[24] invented a new spectrophotometer, and with it he measured the absorption spectrum of visual purple, and related it to vision at low intensities. Koettgen and Abelsdorff (1896),[26] and later Trendelenburg (1904),[34] repeated and extended these measurements; and Garten (1906)[11] added to them sufficiently so that the subject did not die completely. The main current of visual research, however, was concerned with totally different things which bore no relation to photochemistry. That was the great period of color

vision, the time of the stormy debates between Helmholtz and Hering, which kept physiology turbulent for about fifty years, and only quieted down in the early years of the present century.

IV. Indirect Photochemistry

When photochemistry again began to be important in the study of vision, the viewpoint from which the investigation began was neither obvious nor naive, and was completely different from the visual purple days. Photochemistry itself had grown in the intervening period, and had become a science of its own. So that when, twenty years ago, physiology as a whole began to feel the powerful influence of physical chemistry, the drowsing subject of vision shared in the general stimulation. The new interest resuscitated the old photochemical idea of the visual process. However, it was not visual purple which led to the photochemical study of vision, but the reverse. It was the developments in photochemistry which were first applied to visual problems, and these later brought the renewed interest in visual purple which is now current.

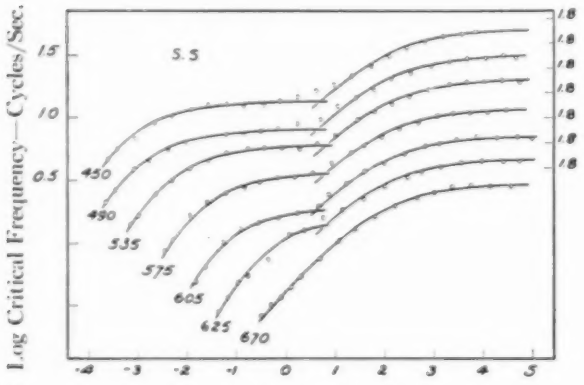


FIG. 2. Critical frequency and intensity for different spectral regions for the eye of S.S. (Hecht and Shlaer, 1936, reference 22) plotted as log frequency against log intensity. The numbers on the ordinates to the left apply to the topmost data; for convenience the others have been moved down in steps of 0.2 log unit, and their exact positions are indicated to the right. The curves are from Eq. (3) in which for the cone portions $m=n=2$, and for the rod portions $m=2$, $n=1$.

The new approach may be formulated somewhat in the following way. Vision in man is undoubtedly a complex series of events. It comprises not only the light receptor processes in the rods and cones but the nerve impulses, which come from them and from nearby places and pass through the optic tracts to be amplified and altered, until they reach the cortex. The visual system of other animals is just as complicated. They all contain receptor systems, nerve impulses, ganglia, and reflexes. All these various elements enter into the end result, and obviously influence its characteristics to some extent. The question is to what extent?

The idea which has dominated the research of the last twenty years, supposes that no matter what enters into the chain of events involving vision, the ultimate place of origin of the impulses which record it is in the initial action of light on the receptor cells. Therefore, in terms of this idea the procedure has been to measure the different functions of vision and photoreception in man and in other animals, and to ascertain how much of their quantitative properties depends on the characteristics of these very first reactions which must take place between light and the sensitive elements. By studying these quantitative characteristics, it became possible to build a hypothetical picture of the intimate chemical and physical events which underlie the first stages of the visual process. The results of this approach have been surprisingly gratifying, and will now be described.

V. Theoretical Formulation

In doing this we may start with the most general ideas about the visual process. There are three things which must take place in the initial events. First, there has to be a sensitive substance which absorbs light and is changed by it into one or more active products. Second, it is necessary to maintain a supply of this photosensitive material; otherwise it would be used up, and the process would come to an end. These two processes may be called the primary light and primary dark

* The unit used here has no relation to the term which the physicist uses for light particles. As used in this paper the term has the meaning first suggested by Troland and is in common use among biologists. It is the retinal illumination produced when the eye looks through a pupil of 1 mm² on a surface whose brightness is 1 candle/m². It is very unfortunate that this duplication of terms exists, and it is hoped that a new term may be chosen in the near future.—Editor

reactions. Third, the active photoproducts of the primary light reaction must do something of which the end result is an impulse from the receptor cell. This is the secondary dark reaction.

These three requirements are the minimum essentials. The photoreceptor process itself is certainly more elaborate. Moreover, vision involves more than the photoreceptor process alone. Yet it is a fact of pertinent interest that the behavior of many visual functions may be formulated in quantitative detail in terms of these minimum essentials alone.

Let us examine the properties of the simplest photochemical system which can be suggested as corresponding to the primary light and dark reactions. Consider a photosensitive substance S whose total initial concentration is a. Let light of intensity I shine on it, and as a result let there be produced the photoproducts P, A, etc., whose concentrations at the moment t are given by x. The rate at which this primary light reaction $S \rightarrow P+A$ goes will be proportional to the light intensity and to the concentration $(a-x)$ of sensitive material; this may be written

$$dx/dt = k_1 I(a-x)^m, \qquad (1)$$

where m is the order of the reaction, and k_1 is a velocity constant which includes the absorption coefficient.

Let us further suppose that some of these photoproducts P, A, etc., can reunite by themselves or with the help of additional substances or energy to form again the sensitive material S. The velocity of this primary dark reaction $S \leftarrow P+A$ will be proportional purely to the concentrations x of the photoproducts; this becomes

$$-dx/dt = k_2 x^n, \qquad (2)$$

where n is the order of the reaction and k_2 is its velocity constant.

When such a system of two reactions is steadily exposed to light, corresponding to the adaptation of the eye to a given light intensity, the two reactions proceed simultaneously until their velocities become equal, and the system reaches a stationary state. Putting Eqs. (1) and (2) equal to each other, and rearranging terms, we get

$$KI = x^n/(a-x)^m, \qquad (3)$$

where $K = k_1/k_2$. This is the equation of the photostationary state of the simple reversible photochemical system $S \rightleftharpoons P+A$.

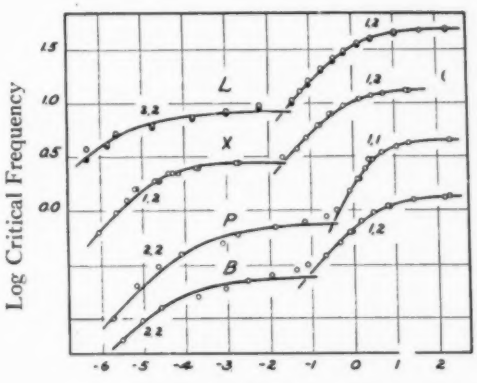


FIG. 3. The relation between critical fusion frequency and intensity for four species of fish. *L* is for the sunfish *Lepomis* (Wolf and Zerrahn-Wolf, 1936, reference 37); *X* is for the swordtail, *Xiphophorus;* *P* for the black platy, *Platypoecilius;* and *B* for the Black Helleri which is a complicated cross between the swordtail and the black platy (Crozier, Wolf and Zerrahn-Wolf, 1937, reference 9). The ordinates apply to the *Lepomis* data; the other measurements have been lowered 0.5, 1.0, and 1.5 log units, respectively, merely to keep them apart. The curves through the measurements are from Eq. (3) with the values of m and n as given for each species.

It is a significant fact that many visual functions behave as if they were controlled by just such a simple system as represented by the stationary state Eq. (3). Indeed, so close is the relation between equation and data, that it is possible to determine from the measurements the numerical values of m and n.

Equation (3) may be studied by putting a equal to 100 percent, and computing the values of KI for different values of x after m and n have been given values of 1 or 2. The results of such a series of computations are shown in Fig. 1. Because of the logarithmic plot, the curves keep their shapes regardless of the magnitudes of K or of the units in which a and x are measured. It is the values of m and n which determine the specific shapes and slopes of the curves.

VI. Visual Investigation

Let us consider some of the visual functions whose properties may be described quantitatively by the stationary state equation. One such function is the perception of flicker when a rhythmically interrupted illumination is viewed

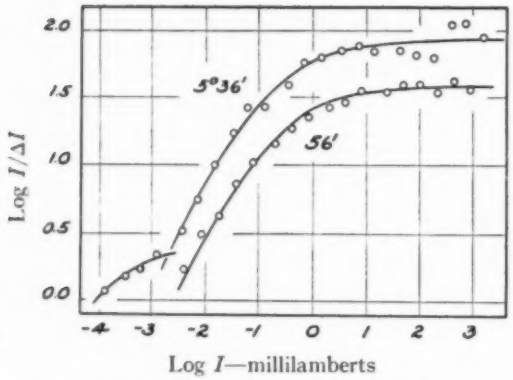


FIG. 4. The relation between intensity discrimination and intensity with different retinal areas and with white light (Steinhardt, 1936, reference 33). The lower data for an area approximately 1° in diameter show only cone function and are described by Eq. (3) where $m=n=2$. The upper data with a much larger area show both rod and cone function. Here too the cone section is described by Eq. (3) with $m=n=2$; while the rod data are too few for critical determination, the curve through them has the same values of m and n as the cone section.

by the eye. The precise data concern the relation between the prevailing light intensity and the critical frequency of interruptions at which the fusion of visual impressions occurs and flicker disappears. Fig. 2 illustrates the relationship for the human eye (Hecht and Shlaer, 1936[22]) between the critical fusion frequency and light intensity for different parts of the spectrum. Because of the different sensibilities of the rods and cones in the spectrum, the relationship appears a double one for all but red light, indicating a separate function for each system of elements.

In terms of fairly simple considerations, it can be shown that in a photochemical system like the one here presented, the critical fusion frequency of flicker at a given light intensity I should be directly proportional to the concentration x at the stationary state. The measurements, of which there are many available, agree with this prediction. In Fig. 2, for each retinal system, the line through the data is taken from Eq. (3) where for the cones $m=2$ and $n=2$, while for the rods $m=2$ and $n=1$. Recently measurements by Crozier, Wolf, and Zerrahn-Wolf (1937)[9] have appeared recording this flicker relationship for a series of fresh-water fish. Fig. 3 shows these measurements and shows also that they may be described with precision by Eq. (3) with various values of m and n, as indicated in Fig. 3.

Exactly the same is true for the intensity discrimination functions of the eye. If I and $I+\Delta I$ are two light intensities which can just be recognized by the eye as differing in brightness, then the fraction $I/\Delta I$ is a direct measure of intensity discrimination. It is simple to show (Hecht, 1935; 1937)[16, 17] that in the hypothetical photochemical system here suggested the fraction $I/\Delta I$ is directly proportional to x^n.

Intensity discrimination has been measured by Aubert (1865),[3] by Koenig and Brodhun (1888),[25] by Blanchard (1918),[4] and by others (cf. Hecht, 1935[16]). All the critical measurements may be described by Eq. (3), in which m and n have appropriate integral values. Steinhardt's (1936)[33] data are in Fig. 4, and our own more recent data (Hecht, Peskin and Patt, 1937[21]) are in Fig. 5. The double nature of the data is again due to the fact that the human eye is a double sense organ of which the rods and cones function independently over different intensity ranges and in different parts of the spectrum (Schultze, 1866;[31] von Kries, 1929[27]).

Still a third visual function, namely visual acuity, falls into this simple agreement with photochemical theory. Shlaer's (1937)[32] measurements, shown in Fig. 6, illustrate the relationship between visual acuity and light intensity. Here, too, we see two functions, except where the measurements were deliberately confined to the rod-free central area of the eye. In all cases, however, the curves drawn through the data are again taken from Eq. (3) with the appropriate values of m and n, since it may be shown that visual acuity is directly proportional to x^n in that equation.

A similar application of these ideas and equations has been made (Hecht, 1937*a*)[18] to the measurements of instantaneous threshold as determined by Blanchard, with the result that they too may be described in terms of the simple photochemical system.

In all these measurements, the recognition of visual function is made by the animal as a whole. For example, when the organism is man, he says that he sees the difference between I and $I+\Delta I$, or that he sees no flicker in the visual field at a specific frequency of interruptions of

the light. When the organism is one that cannot speak, it nevertheless records the visual act by some other characteristic action such as a siphon retraction, or a movement of its body in the direction of a set of moving stripes. A subtle variation of this indirect approach is the use not of the animal as a whole, but only of its eye. In this case the visual effect is recorded by an electrical sign either in the retina or in the optic nerve.

The electrical response of the retina first became known about the same time as visual purple (Holmgren, 1866[23]; Dewar and McKendrick, 1876[10]). However, its use as an indicator of the quantitative relations between light and the receptor system began only in the present century with de Haas (1903)[12] and has come into its own with the work of Adrian (1928)[1] and of Hartline (1935).[13] In addition Adrian's discovery of single nerve impulses has added another sign which has been useful for judging visual function. By these methods Adrian and Matthews (1927)[2] and especially Hartline (for a summary, see Hartline, 1935[13]) have been able to show that the relation between light and the photoreceptor involves a photochemical system having much the same properties as was found by the work with the animal as a whole.

It is necessary again to state that the picture of the initial stages in photoreception, which I have developed here, is probably too simple. Certainly to understand the data of dark adaptation requires the photochemical system to be more than a completely reversible one. Rather it must be a cyclic reaction which is pseudo-reversible, and it may even involve an equilibrium with three different reactions instead of with two. The latter has been suggested by Wald (1935),[35] and is very likely true. The main point of the approach which I have illustrated is that even in terms of the simplest photochemical and chemical notions, it is still possible to describe with excellent precision the quantitative properties of several major visual functions.

This type of investigation of the photochemistry of vision is theoretical and indirect, and has perhaps done its best work. Many of the visual functions have now been described definitively, and additional data will probably not change the description essentially. In all likelihood progress will once again come from the direct photochemical approach. It may be interesting to record that fifteen years ago I already felt the need for the direct investigation of visual purple and actually made some studies of its bleaching in relation to time, light intensity, and temperature. However, I was quickly brought to a standstill by the inadequacies of the times: photoelectric spectrophotometers were not dreamed of then, and the relevant chemical knowledge of proteins and of carotenoids was not available. It took ten years longer before the direct chemical attack could be adequately made.

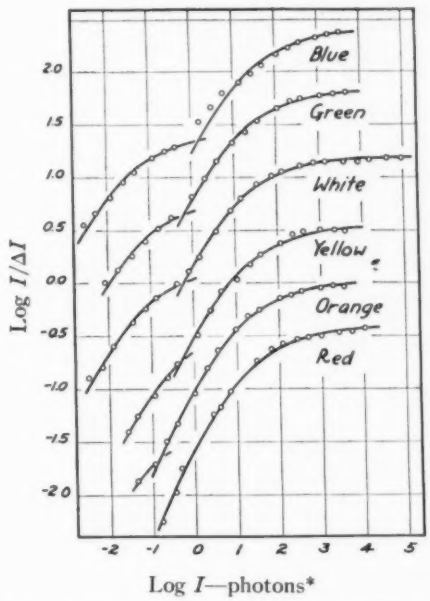


FIG. 5. Intensity discrimination at different intensities for different parts of the spectrum. The ordinates are correct for the measurements with white light; those for green and blue have been raised 0.5 and 1.0 log units, respectively, while those for yellow, orange, and red have been lowered 0.5, 1.0, and 1.5 log units, respectively, merely to separate the points. For red the data are continuous and represent one functional system only, the cones. For the other colors a second system, the rods, becomes evident at low intensities as the color moves toward the blue. As usual in visual measurements, white light behaves much as yellow light. The curves are from Eq. (3) with $m=n=2$ for the cone section; the rod section is also best described by the same values. The discrepancy near the transition between rods and cones for blue light shows up frequently with blue light measurements of other visual functions; *cf.* Fig. 2 for example.

* See footnote on page 158.

VII. Visual Purple Again

What do we now know about the actual photochemistry and chemistry of the substances in the eye? From the physical chemical point of view, we may say very simply that the bleaching kinetics of visual purple are much as one would expect. In a rough way the bleaching follows a first order kinetics (Hecht, 1920*a*);[14] and when care is taken to exclude the additional absorption by new substances formed during the bleaching, the first-order reaction is indeed an adequate description (Chase, 1936).[6] Moreover, the effect of temperature on the bleaching of visual purple is very slight, as is to be expected in terms of photochemical knowledge (Hecht, 1920*b*).[15] More interesting, perhaps, is the fact that the effect of the bleaching light is directly proportional to its intensity, and this too is expected. Spectral absorption measurements made with the new photoelectric methods have confirmed Kühne's discovery that in the bleaching of visual purple a new pigment, visual yellow, is formed, which then changes to visual white. Visual white turns out to be an acid-base indicator, as Chase (1936)[6] has shown; it is yellow in acid solution and almost colorless in alkaline solution. There are many details in the absorption spectrum that are still not clear. Certain it is that the absorption spectrum of the very best preparations of visual purple do not quite coincide with the spectral sensibility of visual purple to light and with its difference spectrum (Lythgoe, 1937;[29] Chase and Haig, 1937[8]). However, these are minor matters and will undoubtedly yield to further research.

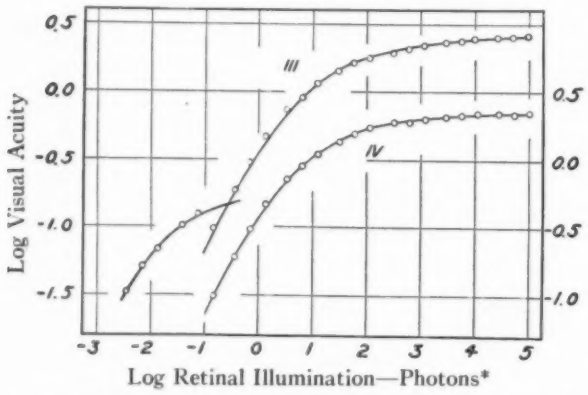


FIG. 6. Visual acuity as related to intensity for two observers (Shlaer, 1937, reference 32). In each case the cone measurements were specifically confined to the fovea. The curves through the data are from Eq. (3) with $m=n=2$ for the cones, and $m=2$, $n=1$ for the low intensity rod section.

* See footnote on page 158.

An arresting application of absorption spectrum measurements to the photochemistry of vision has recently been made by Wald (1937),[36] who was able to demonstrate the presence in chick retinas of two photosensitive materials: one is clearly visual purple; the other, however, is the long suspected, but never previously found, visual violet whose absorption spectrum approximately corresponds to the sensibility distribution of the cones in the spectrum.

From the organic chemical or structural point of view, much has been learned about visual purple in recent years. Kühne could never separate visual purple from the proteins in solution with it. His studies on the destruction of visual purple by heat, already referred to, showed definitely that visual purple is associated intimately with a protein or is itself a protein which is denatured by heat. We (Hecht, Chase and Shlaer, 1937[19]) have studied the diffusion of visual purple through porous glass discs according to Northrop and Anson's method, and have been able to show that the rate at which the color diffuses proves that the molecular weight of the diffusing substance, that is of the color itself, is of the order of 800,000. This renders the protein nature of visual purple almost certain.

Visual purple is not a simple protein however. Wald has found that visual purple in the retina, when bleached, yields retinene, a carotenoid closely related to vitamin A, and has in this way made rational the relation between states of nutrition and vision which has been known for centuries. Visual purple is thus probably a conjugated protein containing a protein group and a carotenoid group. Many such carotenoid proteins are known as animal pigments, though none other than visual purple seems to be sensitive to light.

Sixty years ago Kühne recorded that visual purple even in solution could regenerate after bleaching by light. This observation was never confirmed, though it is an open secret that most investigators concerned with visual purple have vainly tried to repeat it. Recently we (Hecht,

Chase, Shlaer and Haig, 1936[20]) were fortunate enough to discover some of the conditions required for this regeneration to take place in solution. We have been able not only to confirm Kühne's observation but also to measure the course of the regeneration reaction and to determine the absorption spectrum of the regenerated material. In addition, Chase (1937)[7] has found that for visual purple to regenerate in solution another photosensitive substance, present with it, must also be bleached. This new photosensitive material is particularly sensitive to blue light and is different from visual purple, though it may very well be a product of its bleaching.

We are thus at the point where a historical cycle has closed on itself, a cycle lasting only sixty years. The researches which Boll and Kühne began sixty years ago on the direct photochemical investigation of retinal substances had to be abandoned because our knowledge of physics, chemistry, and even of physiology was inadequate. In the intervening sixty years organic chemistry, physical chemistry, and photochemistry have made tremendous strides forward and have supplied us with tools and knowledge for further investigation. Moreover, in the intervening sixty years, the data of vision themselves have been definitively determined and have achieved a form and unity which they did not have in the last century. The near future should therefore be fruitful for our understanding of the intimate nature of the visual process.

Bibliography

1. E. D. Adrian, *The Basis of Sensation: The Action of the Sense Organs* (Christophers, London; Norton, New York, 1928), pp. 122.
2. E. D. Adrian and R. Matthews, "The Action of Light on the Eye: II. Processes Involved in Retinal Excitation," J. Physiol. **64**, 279 (1927).
3. H. Aubert, *Physiologie der Netzhaut* (E. Morgenstern, Breslau, 1865), pp. 394.
4. J. Blanchard, "The Brightness Sensibility of the Retina," Phys. Rev. **11**, 81 (1918).
5. F. Boll, *Zur Anatomie und Physiologie der Retina* (Monatsber. Berlin Akad. 1876), pp. 783.
6. A. M. Chase, "Anomalies in the Absorption Spectrum and Bleaching Kinetics of Visual Purple," J. Gen. Physiol. **19**, 577 (1936).
7. A. M. Chase, "An Accessory Photosensitive Substance in Visual Purple Regeneration," Science, **85**, 484 (1937).
8. A. M. Chase and C. Haig, "The Absorption Spectrum of Visual Purple," J. Gen. Physiol. in press.
9. W. J. Crozier, E. Wolf and G. Zerrahn-Wolf, "Critical Illumination and Flicker Frequency in Related Fishes," J. Gen. Physiol., **21**, 17 (1937).
10. J. Dewar and J. G. McKendrick, "On the Physiological Action of Light," Trans. Roy. Soc. Edinburgh **27**, 141 (1876).
11. S. Garten, Uber die Veränderungen des Sehpurpurs durch Licht," Arch. Ophth., Leipsic **63**, 1 (1906).
12. H. K. de Haas, "Lichtprikkels en Retinastroomen in hun quantitatief Verband," Dissertation, Leiden, (1903).
13. H. K. Hartline, "The Discharge of Nerve Impulses from the Single Visual Sense Cell," Cold Spring Harbor Symposia on Quantitative Biology **3**, 245 (1935).
14. S. Hecht, "Photochemistry of Visual Purple. I. The Kinetics of the Decomposition of Visual Purple by Light," J. Gen. Physiol. **3**, 1 (1920*a*).
15. S. Hecht, "Photochemistry of Visual Purple. II. The Effect of Temperature on the Bleaching of Visual Purple by Light," J. Gen. Physiol. **3**, 285 (1920*b*).
16. S. Hecht, "A Theory of Visual Intensity Discrimination," J. Gen. Physiol. **18**, 767 (1935).
17. S. Hecht, "Rods, Cones, and the Chemical Basis of Vision," Physiol. Rev. **17**, 239 (1937).
18. S. Hecht, "The Instantaneous Visual Threshold After Light Adaptation," Proc. Nat. Acad. Sc. **23**, 227 (1937*a*).
19. S. Hecht, A. M. Chase and S. Shlaer, "The Diffusion Coefficient and Molecular Size of Visual Purple.' Science **85**, 567 (1937).
20. S. Hecht, A. M. Chase, S. Shlaer and C. Haig, "The Regeneration of Visual Purple in Solution," Science **84**, 331 (1936).
21. S. Hecht, J. Peskin and M. Patt, "Visual Intensity Discrimination in Different Parts of the Spectrum," Am. J. Physiol. **119**, 330 (1937).
22. S. Hecht and S. Shlaer, "Intermittent Stimulation by Light. V. The Relation Between Intensity and Critical Frequency for Different Parts of the Spectrum," J. Gen. Physiol. **19**, 965 (1936).
23. F. Holmgren, "Method ot objektivera effekten af ljusintryck po retina," Upsala Läkaref. Förh. **1**, 177 (1866).
24. A. Koenig, "Über den menschlichen Sehpurpur und seine Bedeuting für das Sehen," Sitzungsber. Akad. Wissensch. Berlin, 577 (1894).
25. A. Koenig and E. Brodhun, "Experimentelle Untersuchungen über die psychophysische Fundamentalformel in Bezug auf den Gesichtsinn," Sitzungsber. Akad. Wissensch., Berlin, 917 (1888).
26. E. Koettgen and G. Abelsdorff, "Absorption und Zersetzung des Sehpurpurs bei den Wirbeltieren," Z. Psychol. u. Physiol. Sinnesorgane, **12**, 161 (1896).
27. J. von Kries, "Zur Theorie des Tages- und Dämmerungssehens," in A. Bethe, G. von Bergmann, G. Embden, A. Ellinger, *Handbuch der normalen und*

pathologischen Physiologie, Vol. 13, Part 1 (J. Springer, Berlin, 1929), p. 678.
28. W. Kühne, "Chemische Vorgänge in der Netzhaut," in L. Hermann, *Handbuch der Physiologie*, Vol. 3, Part 1 (Vogel, Leipzig, 1879), p. 235.
29. R. J. Lythgoe, "The Absorption Spectra of Visual Purple and of Indicator Yellow," J. Physiol. **89**, 331 (1937).
30. L. Moser, "Ueber den Process des Sehens und die Wirkung des Lichtes auf alle Körper," Ann. Phys. u. Chem. **106**, 177 (1842).
31. M. Schultze, "Zur Anatomie und Physiologie der Retina," Arch. mikr. Anat. **2**, 175 (1866).
32. S. Shlaer, "The Relation Between Visual Acuity and Illumination," J. Gen. Physiol. **22**, 165 (1937).
33. J. Steinhardt, "Intensity Discrimination in the Human Eye. I. The Relation of $\Delta I/I$ to Intensity," J. Gen. Physiol. **20**, 185 (1936).
34. W. Trendelenburg, "Quantitative Untersuchungen über die Bleichung des Sehpurpurs in monochromatischem Licht," Zeits f. Psychol. u. Physiol. Sinnesorgane **37**, 1 (1904).
35. G. Wald, "Carotenoids and the Visual Cycle," J. Gen. Physiol. **19**, 351 (1935).
36. G. Wald, "Photo-Labile Pigments of the Chicken Retina," Nature **140**, 545 (1937).
37. E. Wolf and G. Zerrahn-Wolf, "Threshold Intensity of Illumination and Flicker Frequency for the Eye of the Sunfish, J. Gen. Physiol **19**, 495 (1936).

And so leaving in your minds the essential role of pure science in modern technological progress and the equal importance of modern technology for the advance of science, I shall close, acknowledging my obligation to these modern philosophers by quoting from one who was a contemporary of Galileo—for it was Sir Francis Bacon who said, "Nature to be commanded must first be obeyed. Knowledge is power."

ERNEST O. LAWRENCE

The Mechanics of Muscular Contraction in Man*

BY W. O. FENN

Department of Physiology, University of Rochester, School of Medicine and Dentistry, Rochester, New York

OF recent years physiologists have been so much concerned with the effort to delve deeper into the actual mechanism of a muscular contraction that they have largely lost sight of other interesting mechanical problems relating to the utilization of the muscle force after it has been developed. Moreover, what little we have taught concerning the mechanics of muscles has largely been exemplified in isolated frog muscles rather than in human muscles where the same phenomena can often be demonstrated equally well or better. It is my purpose therefore, to discuss some of these mechanical aspects of muscular movements in man in the hope of reviving some interest in this neglected subject. It should not be left solely to the attentions of the athletic director and the orthopedic surgeon.

Force vs. Length in Human Muscle

Perhaps the most fundamental mechanical information concerning a muscle is given by the length-tension diagram in which the maximum tension exerted during contraction is plotted against the length of the muscle. The salient feature of such a diagram is the increase of tension with increase of length. This fact is important because it sets up a simple requirement which any theory of muscle contraction must meet—does the theory explain this fundamental peculiarity?

Many such diagrams have been published for frog muscles but not so many for mammalian muscles. In human muscle the diagram presents certain difficulties and requires many observations. Theoretically it may be measured in the flexors of the elbow for example, by a simple spring balance, if the flexor force F exerted near the hand is measured in a direction parallel to the pull of the muscle M. Since the ratio of the lever arms of M and F remains approximately constant at all angles of the elbow, the force measured should increase with extension of the elbow in proportion to the increase of the true muscle force. This method presents certain practical difficulties due to the changing angle of pull at the hand, so it is necessary in practice to measure the force with the spring balance held always perpendicular to the forearm. Such figures are given by Reijs[19] and are plotted in Fig. 1. The force passes through a maximum at about 90° where the lever arm is greatest. Although the exact distance from the axis of rotation at which these measurements were made is not clear in the original report, the figures may be used to calculate the true muscle force in arbitrary units by dividing them by the effective lever arm, b, of the muscles as given by Braune and Fischer[12] for the corresponding angles. The muscle force so calculated increases with increasing muscle length as in frog muscle. Some essentially similar curves for human muscle are given by Franke.[14] It is apparent that as the muscle shortens its diminishing force is in some measure compensated by the increase in its lever arm.

A much simpler way to demonstrate in man the increase of force with increased length of muscle —and one suitable for class purposes—is afforded by the two-joint muscles of the legs, the rectus femoris and the biceps femoris, each of which spans both the hip and the knee joints on the ventral and dorsal sides respectively. The subject lies on his back on a table with the knee flexed at 90° over the edge of the table while the extensor force at the knee is measured. He then sits up, leaning well forward but without otherwise

*Presented at the Symposium on Biophysics, Philadelphia, Pennsylvania, November 4, 5 and 6, 1937.

moving and the measurement is repeated. The one-joint muscles at the knee have not changed their length or position, but the rectus femoris is now much shorter and the total force exerted is correspondingly less. The experiment may be repeated measuring the flexor force at the knee in which case the biceps femoris exerts less force in the supine position because its length is only 42 cm instead of 53 cm[13] as in the sitting-up position. Typical graphical records of such an experiment are shown in Fig. 2. In addition to their convenience for this purpose, the two-joint muscles present some interesting mechanical problems which will be referred to later.

Accustomed as we are to the magnitude of the forces which we can exert at the extremities of our limbs we fail to realize what relatively enormous forces we really have at our command. Thus if the flexors of the elbow can pull 50 pounds against a spring balance held in the hand, the muscles themselves are pulling some 500 pounds. In lifting the body the Achilles tendon has a mechanical disadvantage of about 3. Thus if a 150-pound man can stand on one foot with another 150-pound man on his back and raise himself up on his toes, his Achilles tendon must be pulling 900 pounds or nearly half a ton.

The force which a muscle can exert depends in part upon the "quality" of the individual muscle fiber, the way in which it is stimulated (frequency

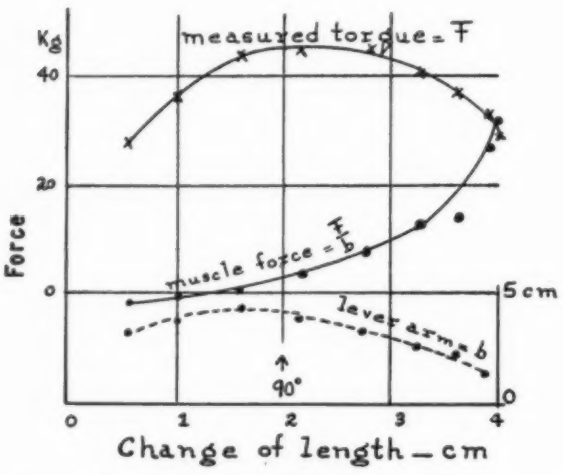


FIG. 1. The muscle force at 90° (F) in kg and the muscle force in arbitrary units (F/b) after correcting for the lever arm. Abscissae represent changes of length calculated from the cosine of the angle and assuming a lever arm of 2 cm at 90°.

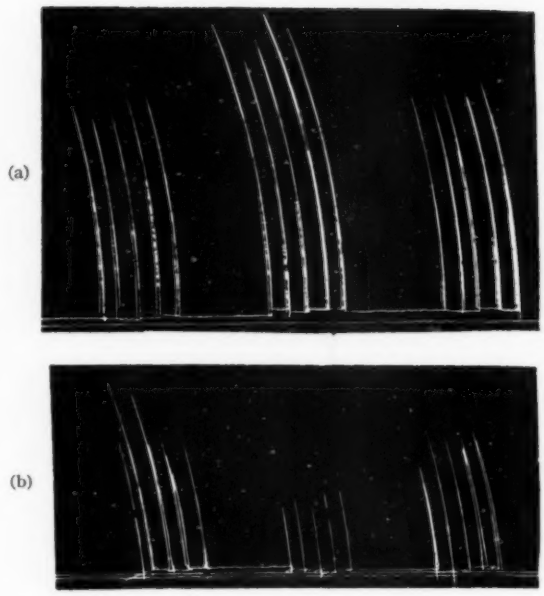


FIG. 2. Extensor (a) and flexor (b) forces of the knee at 90° measured with the hip flexed (sit) and with the hip extended (lie). The force is greater where the angle of the hip is such as to lengthen the two-joint muscles concerned.

of nerve impulses) and upon the arrangement of the fibers in the muscle. The long narrow muscle like the sartorius composed of many units in series provides a small force with large shortening while the broad short muscle, like the glutaeus (upon which we sit) composed of many units in parallel, gives a large force with a small shortening. The force of the glutaeus has been estimated by Fick (reference 12, p. 503) at 1450 pounds. To obtain the convenient shape of the former, coupled with the mechanical characteristics of the latter, the many fibers are often arranged diagonally as in the gastrocnemius muscle between a central tendon and an enveloping sheath. That the arrangement of the fibers is not the only factor which determines the force is shown by the reported performance of a chimpanzee who grasped a rope attached to a spring balance which was passed into her cage and without apparent effort registered a pull of 1260 pounds, an amount 3 to 4 times as great as a human athlete of like body weight could produce, and even greater than the cross section area of the muscles would predict.[2] Similarly, the force exerted by a single frog muscle fiber is far less than that exerted by a human muscle fiber of the same cross section.

According to Franke[14] human muscle can develop 11.1 kg per cm^2 while frog muscle can only develop 1.6–2.9 kg per cm^2 in the sartorius and 3.1 kg in the gastrocnemius.[10] While the force in general therefore, is proportional to the cross section of the fibers, there is evidence that the quality of the fibers varies widely in different species.

Muscle Force at Changing Lengths

Muscles may exert their force in three different ways. (1) Without change of length (isometric). (2) While shortening. (3) While lengthening. "Contraction" of a muscle while lengthening is no more unusual in our bodily movements than the more familiar contraction while shortening. It occurs in all quick movements, the extensors "contracting" while lengthening in order to stop a movement of flexion and *vice versa*.

Relative to these three types of contraction, a problem of unusual interest presents itself for it has been found that the energy expenditure for the contraction is greater than the isometric while shortening, and less than the isometric while lengthening. It is as if the reaction supplying energy for contraction were driven backwards by stretching and forwards by shortening.[8]

Furthermore, in spite of the greater energy production while shortening, the tension produced is diminished as A. V. Hill[16] in particular, has emphasized. The theoretical maximum work or the area under the length-tension diagram is never realized as work in rapid movements. Conversely the tension is greater during a stretch —at least a rapid one—in spite of the diminished energy. The work factor therefore may be more important than the tension in determining the energy liberated. The positive work during shortening demands more energy while the negative work during lengthening demands less energy in spite of converse demands for tension. Evidently there is some thermodynamic regularity about the muscle but the real reason for it awaits a better knowledge of the mechanism of contraction and may yet point the way to crucial discoveries in this direction.

A good example of this contrast between positive and negative work is seen in the problem of jumping up and of jumping down, first discussed by Bethe.[4] The maximum vertical force which the two legs can exert in the various positions taken in arising from a squatting position may conveniently be measured by spring balances stretched between hooks on the floor and a belt fastened about the waist. Plotting this force against the height of the belt above the ground a length-tension diagram is obtained the area of which $\int Tdl$ gives the maximum work which could be done in jumping up from a squatting position. According to Bethe the actual rise h, of the center of gravity of the body (weight $= W$) in such a jump is such that Wh varies from 63 percent to 92 percent of $\int Tdl$. When however, the experiment is reversed and the subject jumps *down* from a height h' such that he can just manage to catch himself in a squatting position without sitting on his heels $Wh > \int Tdl$. The muscles absorb more negative work as potential energy and heat when rapidly stretched than they can produce as positive work when they shorten over the same range of lengths. The differences according to Bethe may be surprisingly large—e.g., $h' = 2h$. (One of Bethe's subjects apparently made an *unfreiwillig* jump from the roof of a shed 6 meters high.) Some of this kinetic energy of the falling body may not be absorbed by the muscles but by oscillations within the body due to the impact with the ground. In addition the muscles, like any viscous elastic body exert somewhat more force during a sudden stretch than at constant length.

There is a possibility that there is something more to Bethe's phenomenon of passive force than mere viscous resistance, but a passive force as great as he describes is difficult to demonstrate in human muscle. On the other hand, if a rapid shortening causes a large fall of tension, an equally rapid stretch should perhaps cause a corresponding rise of tension. Such a rise is shown in Levin and Wyman's[18] observation on dogfish muscle and in Beck's[3] measurements on frog gastrocnemius muscles. More exact statements cannot be made until the force has been measured at speeds of stretch which are accurately measured in absolute units and with proper allowance for the acceleration of the limb.

The Reason for Two-Joint Muscles

In connection with this problem of jumping up from a squatting position I should like to

raise an interesting question concerning the *raison d'être* of the two-joint muscles such as those of the leg which have been discussed above. The action of these muscles can best be illustrated by a consideration of the distinction between concurrent and counter-current movements.* In

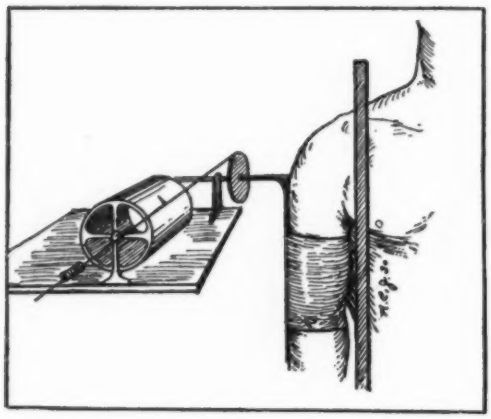


FIG. 3. Apparatus for recording angular displacement of the arm.

the former both hip and knee are flexed together or extended together. Thus the two-joint muscles are being stretched at one end while they are shortening at the other so that their length remains relatively constant, perhaps approaching the isometric contraction. In the counter-current movements the hip is extended while the knee is flexed or *vice versa*, so that the length of the muscle is changing in the same direction at both ends simultaneously, either lengthening or shortening therefore with double rapidity. In running, the last half of the forward movement of the leg is counter-current because the knee is beginning to extend while the hip is still flexing. Here the rate of shortening of the rectus femoris is very high and it must be able to exert only very feeble tension at this point in a fast sprint. During the rest of the running cycle the leg movements are concurrent. In bicycle riding as in squatting on the floor, all the movements are concurrent. Thus in rapid concurrent movements the two-joint muscles offer some real advantage for they permit a muscle to exert tension against a rapidly moving limb without a correspondingly rapid shortening.

There would seem to be, however, some disadvantages to two-joint muscles. Thus a contraction of the rectus femoris to extend the knee will also flex the hip unless this pull is counteracted by some other hip extensor. If the biceps femoris were used for this purpose it would also tend to flex the knee which opposes the movement originally intended. Presumably the glutaeus muscle would be used in this case to extend the hip, in which case this hip extensor might be said to be doing work in extending the knee, its force being transmitted through the rectus femoris muscle. Such an indirect transmission of force, however, carries with it the expense of maintaining tension in the transmitting muscle and represents, therefore, no particular gain.

Another explanation of the two-joint muscles has been offered by Lombard (quoted by Bowen, reference 5, p. 190). According to him, the extensor lever arms of these two-joint muscles are always greater than their corresponding flexor lever arms. This statement seems to be borne out by the figures in Table I, quoted by Fick† (reference 12, p. 501 and 585):

TABLE I. *Maximum change of length due to extension at hip and knee, respectively.*

	HIP cm	KNEE cm	BOTH JOINTS cm
rectus femoris	+6.0	−8.1	−2.1
biceps femoris	−10.0	+5.9	−4.1

In arising from the squat to the erect posture there is therefore a net shortening of both these two joint muscles. The muscle force multiplied by this shortening gives the contribution which these muscles can make to the work of elevating the center of gravity of the body. Since a greater shortening implies a greater lever arm, the simultaneous contraction of both these muscles should cause both hip and knee joints to extend. For both these reasons it would be expected that

* This is Steindler's (reference 22, p. 103) translation of von Baeyer's (reference 1) terms *mitläufige* and *gegenläufige*.

† Fischer (reference 13) gives larger values for the lever arm of the biceps femoris at the hip than at the knee, but according to his diagram the muscle is attached to the ischium so far below the joint that the shortening is less than would otherwise be expected and he finds the muscle length the same in the squatting as in the erect position, i.e., no net shortening occurs in arising from the squat.

these muscles would be useful in arising from a squatting position on the floor. Conversely, they would hardly be of any service to the young lady at the circus who wears suction shoes and walks on the ceiling, should she take it into her head to squat *up* against the ceiling. She would have to rely solely upon one-joint flexors.

Much remains to be learned, I think, concerning these two-joint muscles, why we have them and when they are useful. The latter problem is always complicated by the fact that the muscles chosen for a given limb movement vary widely according to the magnitude and direction of the external resistance to movement which is encountered. Just how the central nervous system manages to determine where a contraction should be requested for every emergency is a baffling question—presumably by long experience guided by proprioceptive trial and error.

Pendulum Movements

Most of our bodily movements are rhythmic to and fro oscillations of the limbs which appear to be sinusoidal in character. The physics and physiology of such pendulum movements repay careful study and involve some of the fundamental principles already discussed—particularly the difference between contraction while shortening, and contraction while lengthening.

Such movements are best recorded by means of an apparatus like that shown in Fig. 3.[9] The axis of rotation of the arm is made to coincide with the axis of rotation of a wheel and the arm is fastened in such a way that the wheel rotates with the arm. The angular rotation of the wheel is recorded linearly on a revolving drum by a pointer pulled by a thread which winds and unwinds from the wheel and is kept taut by a spring. The drum record so taken is a graph of the angular displacement, y, of the arm plotted against time. The slope of this line graphically determined ($dy/dt=\omega$) is the angular velocity. If ω is plotted against t the slope of this second graph is $d_2y/dt^2=A$, the angular acceleration. From A the torque L can be determined if I, the moment of inertia of the arm, is known ($L=IA$.).

Figure 4 represents the analysis of one experiment in which the arm was swung rapidly forward and then back again, all at maximum possible speed and without bending the elbow. The graph shows that the velocity increases to a maximum in the middle of the forward swing and then decreases symmetrically to zero. The back swing shows in the same way a maximum velocity in the middle of the swing. Corresponding to these velocity changes there is a period of forward acceleration, then a period of deceleration which is continuous with a period of backward acceleration and finally another period of deceleration as the arm is brought to rest in its original position.

A point of maximum velocity in the middle of the swing might have been anticipated but it came as a surprise to me at the time as I carried out the experiment. My subjective belief was, that the velocity continued to increase until the arm had reached practically the end of the swing when I stopped it voluntarily and without effort. A little thought would have shown that the same product of torque L and time t necessary to impart momentum $I\omega$ to the arm would likewise be required to take out that momentum and bring the arm to rest. Unless therefore, the strength of the flexors greatly exceeded that of the extensors, the periods of acceleration and deceleration would necessarily be equal. This

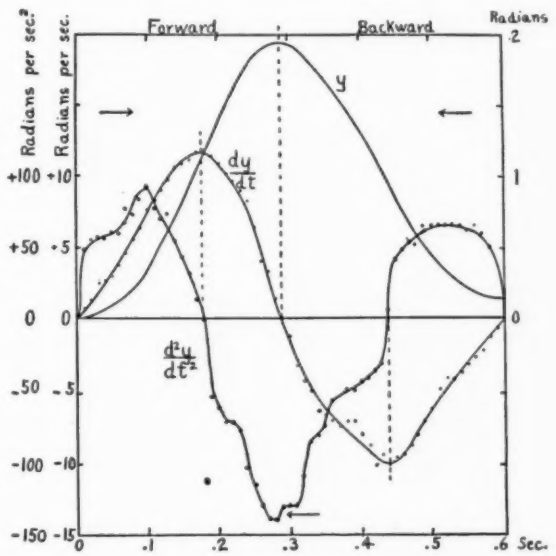


FIG. 4. Graphs from a rapid swing of the arm forward and then back, from Fenn, Brody and Petrilli (reference 9). Velocities and forces in the forward direction are indicated by a plus sign while those in the backward direction are plotted as minus.

subjective error merely goes to show how little conscious concern we have with the minutiae of our bodily movements. It is as if the cortex dictated a rapid movement from back to front and left all the details of the mobilization to the cerebellum.

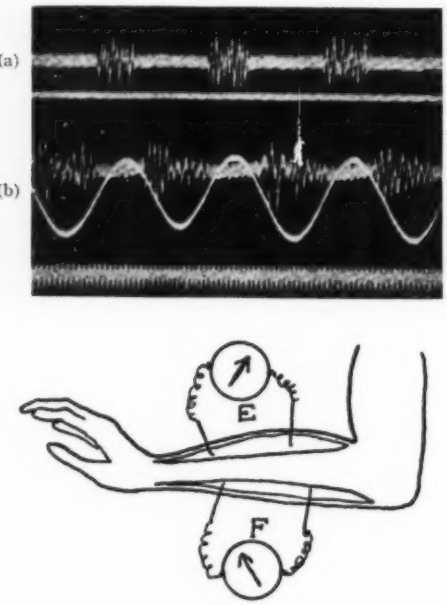


FIG. 5. Simultaneous electromyograms during flexion (down) and extension (up) of wrist from flexor (*a*) and extensor (*b*) muscles (after Wachholder and Altenburger, reference 23) as recorded by 2 string galvanometers. Oscillations represent periods of muscular activity. The sinusoidal curve records the movements of the hand, a down stroke indicating flexion.

More proof of the fact that our bodily movements require contractions of muscles while lengthening, is hardly needed. The above analysis has shown that in the middle of the swing the agonists must relax and the antagonists must go into contraction in order to bring the moving limb to a stop. A convincing confirmation of this fact however, is afforded by the electromyograms of Fig. 5 from Wachholder and Altenburger.[23] Action currents were recorded from both flexors and extensors of the hand during rapid alternate flexion and extension. It is evident that the flexors come into action in the middle of the movement of extension and *vice versa*. There appears to be little or no overlap of the periods of activity of the two sets of antagonistic muscles. This at least happens in fast movements. The slower the movement, however, the less need there is for active contraction of the antagonists until apparently the elasticity of the resting antagonists or the force of gravity is sufficient to bring the moving hand to rest. At this frequency the limb oscillates in its natural period with only a short burst of activity in the flexor muscle at the beginning of flexion and in the extensor muscle at the beginning of extension. By observation of the oscillations of the limb when given a slight displacement, Wachholder and Altenburger[23] estimated the natural period. For the lower arm plus hand and the index finger, they give 1 sec. and $\frac{1}{3}$–$\frac{1}{4}$ sec., respectively. The natural period, t, of these limbs as pendulums can also be calculated from the formula,

$$t=2\pi(k^2/gh)^{\frac{1}{2}},$$

where k is the radius of gyration and h is the distance of the center of gravity from the axis of rotation. Estimating k and h from dimensions of the finger and from data of Braune and Fischer for the lower arm plus hand, I have calculated the periods as 0.57 sec. for the arm and 0.25 sec. for the finger, a fairly close check.

Force in a Moving Limb

As already explained, the velocity and acceleration of the arm can be determined with fair accuracy from the graph showing the displacement of the limb as a function of time. From the acceleration the force can be determined in absolute units if the moment of inertia is known, the equation being $L=IA$ where L is the torque causing the rotation, I is the moment of inertia and A the acceleration. The value of I is arrived at in several ways. A rough estimate may be obtained from the classical data of Braune and Fischer[6] who dismembered a cadaver, suspended the limbs from steel pins in the axes of rotation and calculated the moment of inertia from the period of the swing. An estimate may also be made by familiar formulas from the dimensions of the limb and a knowledge of its density. A more direct method applicable *in vivo* is offered by quick release experiments. The subject exerts a maximal pull against a strong stiff spring which records the force on a smoked drum. The experimenter then pulls a pin suddenly releasing the limb which flies out and records its displacement on another rapidly revolving drum. From one

drum the value of L is obtained and from the other the value of A. Then I, the moment of inertia equals L/A.

By this method the comparisons shown in Table II were obtained which show that the method is capable of considerable accuracy.

TABLE II. *Values of moment of inertia.*

	g cm² × 10⁵ L/A	I FROM BODY WT. ETC. (BRAUNE AND FISCHER[6])
1. Forearm—W.B.L.	0.59	0.557
2. Leg—M.N.	5.35	5.74
3. Leg—W.B.L.	3.77	3.79
4. Leg—A.P.	4.44	4.8
5. Artificial limb and muscles.	2.96	2.99*

* Calculated from familiar formulas for bodies of known shapes.

Having calculated the value of I for the limb in question the acceleration A can be used to calculate the actual muscle torque which is applied to the limb at any moment during a free swing. This can even be done from motion pictures although very limited application has been made of this important possibility. Knowing the force in absolute units some interesting calculations can be made from experiments similar to those in Fig. 4. If the torque is plotted against the displacement in radians and the area under the curve is determined with a planimeter we have the energy used in accelerating the limb to its maximum velocity or in decelerating it. These two values should be equal to each other and they should be equal in turn to the kinetic energy of the limb at its point of maximum velocity as given by the formula $\frac{1}{2}I\omega^2$. Eleven records of a single forcible forward swings of the arm were analyzed from this point of view with the following average results.

Work of acceleration	362 kg cm
Maximum kinetic energy	386
Work of deceleration	365

Considering the amount of graphical analysis involved this agreement is quite satisfactory. As an indication of the striking symmetry of these movements already mentioned, it may be added that the average durations of the period of acceleration and deceleration respectively were 0.187 and 0.186 seconds. The total angle of movement was 105 degrees of which 52° were acceleration and 53° were deceleration.

Of particular interest in connection with these arm swings is the fact that the maximum torque which can be exerted by the flexor muscles at the point of reversal of direction as calculated graphically from the acceleration, was closely equal to the maximum torque which they could exert statically against a dynamometer when they were contracted maximally with the arm in the same position. In a forcible swing of the whole arm therefore, the muscles contract isometrically and with maximum tension at the point of reversal. In vibrations of smaller parts such as the forearm or the fingers the maximum force is not attained.

Force vs. Displacement

Inspection of the graphs in Fig. 4 shows a certain parallelism between the force d^2y/dt^2 and the displacement y. Such a proportionality is expected if the displacement curve is sinusoidal. In a pendulum, however, the force is determined only by the displacement and has a fixed value for every position which is quite independent of the amplitude of the swing. In arm movements the force is roughly proportional to the displacement for any given amplitude of vibration but when this amplitude changes the whole force-displacement pattern changes. Evidently the central nervous system "deals out" excitation to the appropriate muscles according to a sine wave pattern but that pattern varies with the ampli-

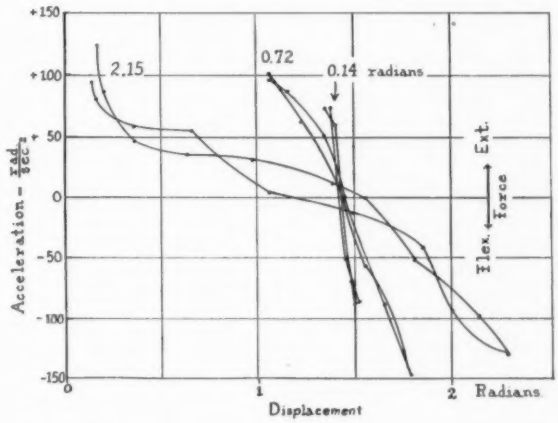


FIG. 6. Data from records of oscillatory movements of the forearm at 3 different amplitudes (as indicated on the graph) and all at maximum speed. The displacement in radians is measured from the point of maximum flexion as an approximate zero. Each graph represents one complete cycle out of a continuing series. The moment of inertia of the arm about the elbow is about 5.6×10^5 g cm.²

tude. Likewise the period varies with the amplitude (unlike a pendulum) although the force at the smaller amplitudes is definitely less than at large amplitudes.

These facts may be illustrated by a series of experiments in which the forearm was vibrated at different amplitudes but at maximum possible velocity. Records were taken by an apparatus similar to that illustrated in Fig. 3, the axis of rotation being made to coincide with the axis of the elbow. The curves so obtained were analyzed according to the method already described. The force exerted at three different amplitudes has been plotted in Fig. 6 as a function of the displacement in radians. The three curves have been arbitrarily placed on the graph so that they cross the zero line at the same point, the correct location being inaccurately known. While the points are not perfect straight lines, there is a rough proportionality between force and displacement. The important point is that the curve obtained at small amplitudes is not merely a small section of that found at large amplitudes, but has instead a totally different slope of its own. Moreover, the maximum force attained in each of the three curves is not proportional to the corresponding amplitude of displacement, but is excessively large at small amplitudes. In general, the maximum force is great at large amplitudes but it may occasionally be larger at intermediate amplitudes than at the greatest amplitude. The maximum possible force exerted in these fore-arm movements is only about one-fourth as great as the force which could be exerted if the limb were not moving. In this respect, the forearm differs from the case of the whole arm swinging from the shoulder, which was illustrated in Fig. 4. Probably, continuous to and fro oscillations develop rather less force than a single swing forward and backward, in which the attention can be more definitely focused on an effort to keep the velocity maximum. Apparently also the smaller the limb, the smaller the force which can be exerted in high speed oscillations. In the fingers the force is even a smaller fraction of the total available force than it is in the forearm.

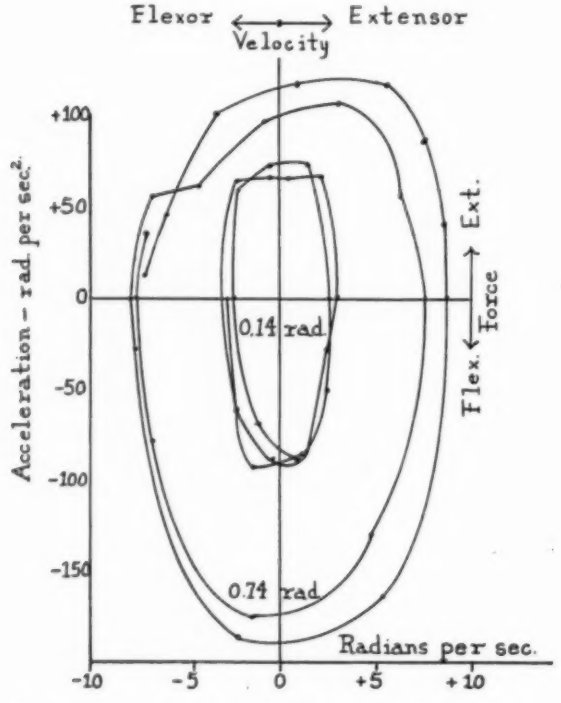


FIG. 7. Successive changes of force and velocity during to and fro movements of the forearm at amplitudes of 0.14 and 0.74 radians. The movement progresses in a clockwise direction. The moment of inertia of the arm is about 5.6×10^5 g cm.²

Force vs. Velocity

It may be suggested that the force exerted in these to and fro oscillations may depend upon the velocity of movement. It is of interest therefore, to plot the force as a function of the velocity. In sinusoidal movements a circle should be obtained if the velocity is plotted against the displacement. Since force is proportional to displacement the force-velocity curves should also be circular if the scale is properly selected. The graphs of Fig. 7 show that this expectation was approximately realized. Data from two different amplitudes are included—the smaller oval representing the smaller amplitude. The movement is in a clock-wise direction, points at top and bottom representing the ends of the swings where the velocity is zero and the points at the sides, right and left, representing the midpoints of the swing where the velocity is maximum and the acceleration zero. In general the velocity like the force is greater at the larger amplitudes. It seems probable that the maximum velocity is limited not by the properties of the muscle itself, but rather by the necessity that the agonists relax in order to permit the antagonists to bring the limb to a standstill.

These data show that in our normal bodily movements the force may be said to vary as a circular function of the velocity. The same is

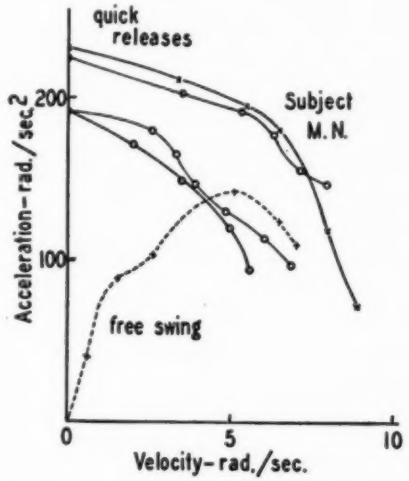


FIG. 8. Decrease of force with increase of velocity in a (lower) leg which is suddenly and unexpectedly released when the extensor muscles are tensed to the maximum (reference 9). The dotted line shows the same movement but without previously tensing the muscles and without restraint.

true in quick-release experiments on the leg similar to those described above for the estimation of the moment of inertia. Here the leg is suddenly released with the muscles previously tensed to a maximum. As the leg of the unsuspecting subject flies out, its acceleration or force decreases as the velocity increases, the force-velocity curve being concave to the origin as in Fig. 7. Some such curves are shown in Fig. 8 together with the dotted line showing how the force increases when the same movement starts without previously contracting the muscles. Actually it is not surprising that the quick-release graph should resemble that found in free to and fro oscillations for at every point in the swing where the direction of movement is reversed the limb is mechanically in exactly the same position as in a quick-release experiment immediately after the release, i.e., the force is already developed in the muscles and the limb is free to move. The only difference is that the force is not quite a maximum in the to and fro movements except in a single swing of the whole arm carried out with greatest effort (Fig. 4).

Other data concerning the force-velocity problem will be considered under the heading of viscosity.

It is remarkable how precisely the arms move in a sine wave fashion, especially when it is realized that the movement is dependent upon the net muscular force exerted and that this force at every point is determined not simply by the length of the muscles as in a spring but by a nicely coordinated nervous system which, even in individuals untrained in physics, seems nevertheless to be cognizant of the mathematical requirements of a sinusoidal oscillation.

Frequency of Vibration

A study of the maximum frequency of oscillation of different parts of the body at minimum amplitude shows that here too, a nervous factor is primarily concerned rather than a purely physical factor. This is especially obvious with the small limbs like the fingers. To show this I determined the maximum frequency of vibration of a leg, an arm, a finger and a toe and compared the period in seconds with the theoretical acceleration which the muscles could produce if maximally contracted.[8d] The acceleration was calculated from the ratio of the maximal flexor torque (as measured statically) to the moment of inertia, i.e., $A = L/I$. The resulting figures are shown in Table III. It is evident that the period

TABLE III.

	PERIOD AT MAX. FREQUENCY SEC.	THEORETICAL ACCELERATION RAD./SEC.²
Whole leg	0.24	71
Lower leg	0.33	298
Forearm	0.15	775
First finger	0.18	42,500
Great toe	0.30	85,800

does not differ much for these different parts of the body but the theoretical acceleration is more than one hundred times greater for fingers and toes than for the larger limbs. Evidently in wagging these parts the full muscular force available can never be utilized (as has already been mentioned) for if it were, the frequency would be much greater; the limiting factor is rather the speed with which excitation and inhibition can be made to alternate in the central nervous system without at the same time losing pre-

cise control of the magnitude of the force. It has of course, long been regarded as a characteristic of the nervous system that a maximum frequency of 6 to 10 per second was observed for such miscellaneous efforts as counting, reciting the alphabet, tapping, wagging the tongue,[17] and muscle tremors.[20]

Frequency vs. Amplitude

The effect of amplitude upon the maximum possible frequency of oscillation yields some figures of interest in this connection. Records on a moving drum were taken by an apparatus similar to that shown in Fig. 3 but modified for use with other limbs. To and fro oscillations of different amplitudes were carried out as rapidly as possible. A few typical experiments are shown in Fig. 9 where the amplitude is plotted against the time for one complete cycle. The graphs obtained are similar in all cases and the period is given approximately by the equation $T=t+bA$ where b is a constant, A the amplitude, and t a minimum time representing the period at the smallest amplitude. That the equation is of rather general applicability to rhythmic movements is shown by the accuracy with which it applies even to the movements involved in writing large and small letters at maximum speed.

The interpretation of this equation is not easy. It is not satisfied by assumptions based on the theory of a pendulum. In a pendulum the period should of course be (nearly) independent of the amplitude, the force being sufficiently larger at the larger displacements to produce the necessary

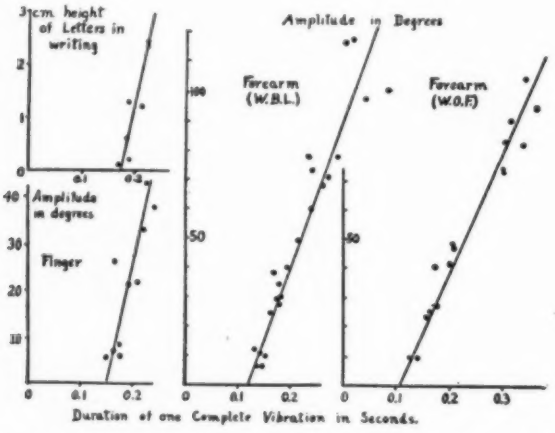


FIG. 9. The relation between the period (abscissae) and the amplitude of vibration (ordinates).

increase in velocity. The relative constancy of the time t, for all limbs both large and small has already been discussed (cf. Table III) and it seems probable that it represents the time which the central nervous system requires to reverse the direction of movement of a limb without disturbing its orderly sinusoidal progression. Apparently movement is never reversed as abruptly as the force available would make possible, but it is done in an orderly manner and under adequate proprioceptive control. If this represents the correct interpretation of t in the equation then the constant b represents an average, more or less constant, velocity of movement, the time bA needed for the observed displacement A being then proportional to the amplitude. The chief difficulty in this interpretation is the fact already mentioned that the maximum velocity is not constant but greater at the larger amplitudes. We must not insist too strongly therefore, upon the use of this particular equation. The data do not actually fit so closely as to suggest that no other equation could be used instead.

It might be suggested that the time t represents the minimum time necessary for the flexor and extensor groups of muscles to contract and relax in sequence. A single twitch of a cat's rectus muscle produces a mechanical response lasting about 0.04 sec. (Reference 15, p. 210.) Alternating but nonoverlapping twitches of the flexor and extensor muscles would therefore give a minimum period of 0.08 sec. Since the minimum periods observed are twice as long as this it appears that single twitches will not suffice even for the minimum amplitude of movement of a finger, or that the central nervous system is unable to deal in stimuli of such meager duration.

In a voluntary contraction of a human arm the period of rise of tension from zero to maximum is about 0.04 seconds. If a similar period sufficed for relaxation the total period for a complete oscillation would be 0.16 seconds. This agrees closely with the observed value of 0.15 sec. for the forearm but the force actually developed for such an oscillation of minimum amplitude is only about one-tenth of the maximum possible. Thus at an amplitude of 0.14 radians illustrated in Fig. 7 the maximum recorded acceleration is about 80 rad./sec.^{2} while the calculated maximum for the author's arm is 775 (Table III).

Much smaller times would be needed for the development of forces which are relatively so small. It would appear therefore, that still higher frequencies of oscillation could be obtained if the appropriate muscles could be stimulated electrically in proper sequence. Again, it is not merely the development of force but the *orderly* development of that force which limits the speed of oscillatory movement.

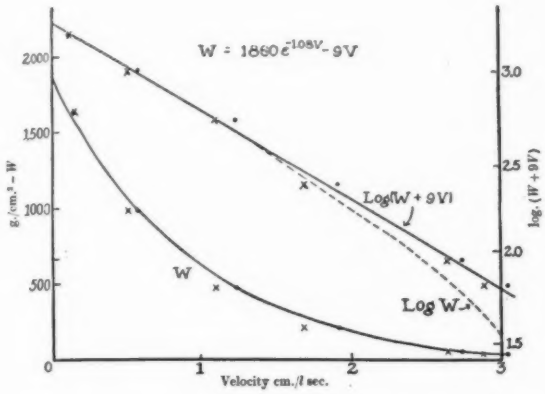


FIG. 10. The decrease of force (W) with increase of velocity of shortening. The lower curve and the straight line (log W+kv) are drawn to fit the equation. (Reference 10.)

Viscosity

The importance of viscosity in limb movements was first suggested by A. V. Hill[16] in connection with the diminished force which can be exerted by muscles at high speeds of shortening. His measurements showed that the decrease of force with increase of velocity is linear. He assumed therefore, that at high speeds of movement the viscous resistance was increased in proportion to the speed so that correspondingly less force could be exerted by the muscle externally. Hill[16a] found an application and confirmation of this theory in his study of the velocity of a runner at different times in a sprint, the observed relation being consistent with the assumption that the resistance to movement is proportional to the speed. At the beginning of the sprint all the force is available for acceleration while at constant velocity it is all used to overcome the viscous resistance. Likewise a linear force-velocity relation was found by Dickinson[7] in bicycle pedaling against different resistances.

While this theory has proved extremely stimulating to workers in this field, certain objections to it can now be offered. In the first place if the efficiency of muscular exercise is the same in running as in other types of muscular work then the mechanical work done should be about 25 percent of the total energy as measured by the oxygen consumed. If all the work of sprint running is done against viscous resistance therefore, there should be no appreciable external work done. The work of running includes the work of accelerating the limbs plus the work of decelerating them plus work against gravity and work against wind resistance. These items were all measured from high speed moving pictures of 21 sprinters[8b, c] and it was found that the external work thus accounted for amounted to 23 percent of the total oxygen.* This suggests therefore, that the viscous resistance is much less than Hill's theory assumed.

Secondly, it has been shown by Fenn and Marsh[10] that the true relation between force and velocity is not a linear but a logarithmic decrease. This conclusion resulted from experiments on cat and frog muscles in which contraction was caused by electrical stimulation of the muscle itself (frogs) or its motor nerve (cats). Thus variations which might be caused by the central nervous system were avoided. During the stimulation the muscles shortened under different loads (W) and the velocity of shortening (v) was measured. Some typical results are shown in Fig. 10. The lower graph shows that W does not decrease *linearly* with increase of v as Hill's original theory would predict. Except at high velocities, however, a close approach to a straight line is obtained by plotting log W against v (dotted line). The deviation from a straight line at high velocities may be attributed to a viscous loss since with suitable choice of a coefficient of viscosity k, a straight line is obtained (upper graph) if log $(W+kv)$ is plotted against v. This relation indicates that the true internal force F exerted by the muscle falls off with increase of velocity

* The total rate of energy turnover in the sprint amounts to 13 horsepower. The total rate of output of measured mechanical work is 2.95 horsepower, of which wind resistance accounts for 0.5, gravity for 0.1 and the increases and decreases of kinetic energy of the limbs for 2.3 horsepower.

according to an exponential law, the force manifested externally (W) being somewhat less at high velocities because of the viscous resistance (kv). Thus the relation between external force, W and velocity v, in a muscle freed from any possible nervous influences was given by the equation,

$$F = W + kv = W_0 e^{-av},$$

where W_0 is the external force exerted isometrically, a is a "coefficient of tension loss" and k is the coefficient of viscosity. While this is a purely empirical equation it does fit the facts as they are now known, not only in muscles of frogs and cats but also in dogfish, scollops, and *Phascolosoma* worms (unpublished work of J. B. Hursh).

Evidently the linear force-velocity relation found by Hill[16] must be given some other interpretation. Perhaps if his observations could have been carried to still higher speeds of shortening the linear relation would no longer have held. The decrease of force attained was never as large as in the experiment of Fig. 10. It is also likely that the voluntary innervation is not the same at high velocities as at low velocities. Finally, Hill's original experiments on arm flexion were concerned only with average velocities of shortening and in his running experiments he used in his equation the velocity of the whole runner while the velocities of shortening of the muscles themselves varied over wide limits from plus to minus values.

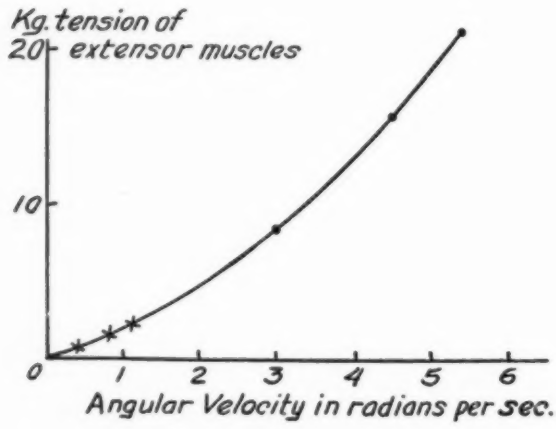


FIG. 11. Resistance of the normal human leg to passive movement about the knee joint (in kg at 3.8 cm from the axis) as ordinates and velocity of movement as abscissae. Dots from the fall of the leg by gravity; crosses from movement of the leg by machinery at constant velocity. (Reference 11.)

Thus the true muscular force-velocity relation is *exponential*, with viscosity a fairly unimportant factor. The force-velocity relations during the course of any one actual limb movement are *circular*. As a combination of these two extremely different effects a linear relation may be found over a limited range where mean forces and velocities are concerned.

Resistance to Passive Movement

It might be suggested that the problem of the role of viscosity could be settled by measurements of the resistance offered by the leg to passive movement at different velocities. Some data of this sort were obtained from two series of measurements of the muscle tonus of some normal and pathological subjects. Two different methods were employed. In the first,[21] the lower leg was simply allowed to fall by flexion of the knee from the horizontal to the vertical position by the force of gravity. A graphical record was taken of the rate of fall; from the record the angular acceleration could be estimated. Except in the most flaccid legs this calculated torque was less than the calculated torque due to gravity by an amount which represented the frictional resistance to movement and the tension of the extensor muscles opposing the movement. From a single fall the magnitude of this resistance could be evaluated for several different velocities at different stages in the fall, and it is certain that the frictional resistance to such passive movement could not be larger than the amount measured.

In another method[11] an apparatus was constructed which permitted the lower leg to be alternately flexed and extended by the power of an electric motor, the whole movement taking place with constant angular velocity except at the ends of each stroke where the direction of the movement was reversed. The force exerted by the machinery against the leg at constant velocity was recorded directly on a revolving drum. This force represents the frictional and elastic resistance to movement and is necessarily quite independent of the inertia of the leg. Elaborate graphical analysis is therefore unnecessary. This force was measured for normal and pathological

subjects at varying velocities. The data yield average figures for the increase of force with increase in the velocity of movement. Fig. 11 shows the points obtained by these two methods. Within limits of the errors involved this curve is not logarithmic as in active muscle, but linear as pure viscosity should be. If the slope of this curve is estimated in terms of gms. per cm.^{2} of muscle for an increase of velocity of 1 cm per cm length of muscle a value of about 283 is obtained (corrected from Fenn and Garvey[11]) which includes elastic resistance as well as friction. A similar value from Hill's experiments on arm flexion is 10,100. Our experiments on quick release of the leg have given values as high as 3000. While the uncertainty of these estimates is relatively enormous, it may perhaps be concluded that resistance to passive movement is not a large factor in accounting for the much greater loss of force with increase of velocity which is found in actively contracting muscles.

Much more could be said concerning the mechanics of muscular movements in man. The literature is full of descriptions of strange experiments on human arms and legs, which have been weighed and jerked, vibrated and released, flexed and extended, at varying rates and amplitudes, in normal and pathological subjects. Problems of locomotion, of running, walking, jumping, flying, throwing and kicking, pushing and pulling, cranking and winding have been studied by all sorts of methods and apparatus. Too frequently this is done with no adequate comprehension of the physics of the process and no attempt to make any accurate measurement of anything in intelligible absolute units. Nevertheless, there is growing up a body of information on the subject. Fundamental problems still await their solution. Some are largely anatomical like those relating to the two-joint muscles; some are partly neurological like the force-velocity problem in voluntary contraction; some are chiefly myological like those relating to force and energy in shortening and lengthening muscles. All are fundamentally physical and will not only yield valuable information but will assuredly grow in importance and breadth, the more thoroughly they are cultivated by competent biophysicists.

Bibliography

1. H. v. Baeyer, Zeits. f. Anat. Physiol. Anat. Hefte, **64**, 276 (1922).
2. J. E. Baumann, J. Mammalogy **7**, 1 (1926).
3. O. Beck, Pflügers Arch. f. d. ges. Physiol. **199**, 481 (1923).
4. A. Bethe, Pflügers Arch. f. d. ges. Physiol. **222**, 334 (1929).
5. W. P. Bowen, *Applied Anatomy and Kinesiology*, fifth edition. Revised by R. L. Boughner and E. J. Rynearson, Philadelphia, 1934.
6. W. Braune and O. Fischer, Abh. d. Math. phys. Kl. d. Sachs. Akad. d. Wiss., **21**, 153 (1894).
7. S. Dickinson, Proc. Roy. Soc. **B103**, 225 (1928).
8. W. O. Fenn, J. Physiol. **58**, 297 (1923). [a] J. Physiol. **58**, 373 (1923). [b] Am. J. Physiol. **92**, 583 (1929). [c] Am. J. Physiol. **93**, 433 (1930). [d] Pflügers Arch. f. d. ges. Physiol. **229**, 354 (1932).
9. W. O. Fenn, H. Brody and A. Petrilli, Am. J. Physiol. **97**, 1 (1931).
10. W. O. Fenn and B. S. Marsh, J. Physiol. **85**, 277 (1935).
11. W. O. Fenn and P. H. Garvey, J. Clin. Investigation **13**, 383 (1934).
12. R. Fick, *Handbuch des Gelenke*, Vol. 3, p. 318.
13. K. Fischer, Zeits. f. Anat. u. Entwicklungsgesch. **83**, 752 (1927).
14. F. Franke, Pflügers Arch. f. d. ges. Physiol. **184**, 300 (1920).
15. J. F. Fulton, *Muscular Contraction*, (Williams and Wilkins Co., Baltimore, 1926).
16. A. V. Hill, J. Physiol. **56**, 19 (1922). [a] *Muscular Movements in Man*, (New York, 1927).
17. C. V. Hudgins and R. H. Stetson, Arch. neerland de phonetique exper. **13**, 85 (1937).
18. A. Levin and J. Wyman, Proc. Roy. Soc. **B101**, 218 (1927).
19. J. H. O. Reijs, Pflügers Arch. **191**, 234 (1921).
20. E. A. Schaefer, E. L. Canney and J. O. Tunstall, J. Physiol. **7**, 111 (1886).
21. A. E. Smith, D. S. Martin, P. H. Garvey and W. O. Fenn, J. Clin. Investigation **8**, 597 (1930).
22. A. Steindler, *Mechanics of Normal and Pathological Locomotion in Man*, (Baltimore, 1935).
23. Wachholder and Altenburger, Arch. für die ges. physiol. **214**, 625 (1926).

Some Aspects of Nuclear Physics of Possible Interest in Biological Work*

BY L. A. DUBRIDGE

University of Rochester
Rochester, New York

IN selecting a title for this paper I have endeavored to choose one which would create no illusions as to its purpose or content. It is my purpose only to review (in a very elementary way) some of the aspects of nuclear physics which give promise of providing useful tools in biological research. I shall attempt to indicate a few of the types of problems to which these tools may be applied. I am not reporting any research which has been carried on in this field. I cannot claim originality for any of the suggestions I am going to make concerning biological problems, since they are all either more or less obvious or else have been proposed by various workers interested in this field.[1] Note that these are only suggestions as to possible problems to which these new tools may be applied; not predictions of any definite results which may be obtained. It is as difficult to make predictions in this field now as it would have been thirty-five years ago to predict the biological applications of x-rays and radioactivity.

Nevertheless, despite these cautious remarks, it is my own conviction, inspired by the conviction of many others better able to judge, that the discoveries of the last five years in nuclear physics are almost certain to be of far reaching importance in biology. They will greatly facilitate experimental work now going on in certain fields, and they will doubtless uncover new problems not now suspected or not open to attack by present experimental methods. Many feel that a new era in biological technique is now at hand. If this be true no further apologies need be offered for any attempt to acquaint biologists with the nature of the new tools for research which they are likely soon to be using.

We will do well to begin the discussion by listing some of the most important advances in nuclear physics in recent years.

1. The discovery, measurement, and separation of isotopes of the chemical elements (Aston, Bainbridge, Urey, *et al.*)
2. The development of machines for accelerating charged particles to very high energies (1 to 8 million electron volts)
   a. Transformer-condenser methods (Cockcroft and Walton, Lauritsen, *et al.*)
   b. Electrostatic generator (Van de Graaff, Tuve)
   c. The cyclotron (Lawrence, Livingston and Cooksey)
   d. Linear accelerators, resonance transformers, etc.
3. The transmutation of elements
   a. By natural alpha-particles (Rutherford, 1919)
   b. By artificially accelerated protons (Cockcroft and Walton, 1932)
   c. By deuterons (Lawrence, 1933)
   d. By neutrons (Feather, 1932)
4. Discovery of the positron (Anderson, 1932)
5. Discovery of the neutron (Chadwick, 1932)
6. Discovery of induced radioactivity (Curie-Joliot, 1934)

We shall pass over the field of isotopes without discussion. The heavy isotopes of hydrogen, carbon, oxygen, etc. are of great importance to nuclear physics and to biology but their use in the latter field is largely a problem of chemistry—and I shall leave it to the chemists to discuss.

* Presented at the Symposium on Biophysics, Philadelphia, Pennsylvania, November 4, 5 and 6, 1937.

Similarly it will not be profitable in this discussion to describe in detail the experimental techniques which have been developed for producing high energy particles. It is the development of these techniques of course which has made possible all the other advances: without them it would not be possible even to discuss in a practical way the possible biological applications of nuclear physics. In fact, there would be very little nuclear physics. The time may not be far distant when a cyclotron or a Van de Graaff generator will be standard equipment in large biological and medical research centers. But our chief interest in this symposium is what these machines will do rather than how they work. The principle thing they will do is to produce transmutations of elements, and the products of these transmutations are the things which biologists may find useful. I should also mention, however, that resonance transformers and Van de Graaff machines are also suitable, and are even now in use in several medical centers, for the production of million-volt x-rays. Still higher voltages can be produced as soon as they are required, but it will be best to understand something of the effects of one-million volt x-rays before two-million volt machines are built. These machines bring x-ray energies right into the realm of gamma-rays, with available intensities which are thousands of times greater than those attainable from natural radioactive sources.

The great importance of extending radiological investigations into this region is sufficiently obvious. There is but one interesting point worth mentioning. One advantage usually ascribed to high energy x-rays is their greater penetrating power. However, this increase of penetrating power with increasing energy (decreasing wavelength) does not continue indefinitely. Above about 2.5 million volts (Mv) the penetrating power of x-rays in lead begins to decrease again, so that 3 Mv rays may actually be "softer" instead of "harder" than 2 Mv. For lighter materials than lead the maximum of penetrating power occurs at rather higher energies (10–12 Mv for Al) but the existence of this maximum must be taken into account, and may offer difficulties in the development of absorption and filtering techniques which will have to be used with these multi-million volt machines.

But million-volt x-rays are really not new tools for biological work—they are only more powerful forms of tools already familiar. We must look to the results of experiments on the transmutation of elements for the really new possibilities which nuclear physics will furnish to biology. Among these the most promising are the last two listed above, namely, the neutron and induced radioactivity, and to these subjects I shall devote the remainder of this discussion.

The Neutron

The neutron is a particle of mass 1 (actually 1.0090) and charge zero. It is one of the two fundamental building stones from which all nuclei are constructed, the other being the proton (mass 1.0076, charge +1). All nuclei are believed to be composed of these two particles and no others. Neutrons may be ejected from various nuclei by bombarding them with protons, deuterons, alpha-particles or gamma-rays. And this, in fact, is the only way in which neutrons can be made available for use in experimental work. When radon (radium emanation) is collected in a small capsule containing powdered beryllium a convenient source of neutrons is

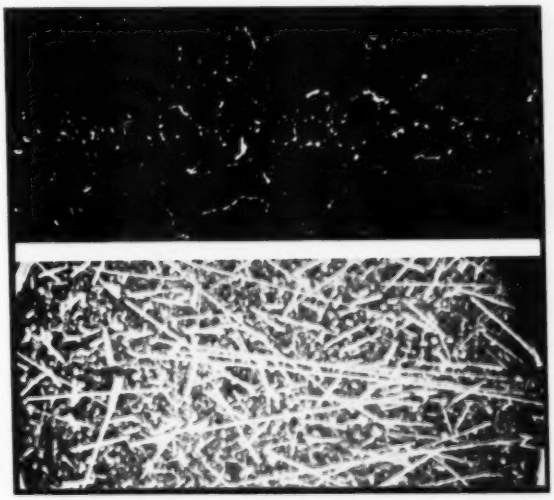


FIG. 1. The difference in biological action between x-rays and neutrons is illustrated above. The upper figure made by C. T. R. Wilson shows the tracts of electrons ejected from atoms in a cloud chamber by x-rays while the lower one made by E. O. Lawrence shows the recoil protons produced by neutrons traversing a cloud chamber.

obtained. The neutrons are ejected by the bombardment of the beryllium nucleus by the alpha-particles from radon according to the reaction

$$Be^9+He^4=C^{12}+n^1.$$

The neutrons emerge with energies up to 13.7 million electron volts (Mev) though most of them have much lower energies.[2]

While the intensity of the neutron beam obtainable from a radon-beryllium source is sufficient for many purposes, enormously more intense beams have been produced by the cyclotron. Lawrence has reported neutron beams equivalent to what would be produced only by several hundred kilograms of radium and beryllium. The reaction often used in this case is the bombardment of beryllium by deuterons, according to the reaction

$$Be^9+H^2=B^{10}+n^1.$$

What then are some of the properties of neutrons so produced? In some ways a beam of neutrons is similar to a beam of gamma-rays. They are undeflected by a magnetic field and have considerable penetrating power through heavy substances. They were, as a matter of fact, mistaken for gamma-rays in the early experiments. A further examination shows however profound differences between neutrons and gamma-rays—and these differences may have biological importance. A gamma-ray loses energy in passing through matter largely through the energy it imparts to the orbital electrons of the atoms through which it passes. A gamma-ray, being an electromagnetic radiation, has strong interaction with charged particles (electrons) and these will sooner or later rob each gamma-ray of all its energy. A neutron, however, being an uncharged particle, is not at all influenced by passing through an atmosphere of electrons. Its course can be deflected or its velocity changed only by a close encounter with the nucleus of an atom, and this is a rather rare occurrence. Even when it does occur, if the nucleus against which the neutron collides is heavy, the neutron will rebound without much loss of energy, like a golf ball from a brick wall. The neutron is scattered but does not slow down.[3] The neutron is completely absorbed only in case it enters the nucleus, and is there captured or causes the ejection of another particle. In most substances this absorption probability is very small and the neutrons therefore penetrate great thickness—e.g., several feet of lead. On the other hand some substances, for example cadmium, have a very large selective absorption for neutrons of a particular velocity—and so rapidly remove neutrons of this particular velocity from the beam.

When neutrons pass through light materials, especially materials containing hydrogen, the situation is quite different. In collision with a hydrogen nucleus, which has about the same mass as a neutron, the neutron may lose a large fraction or all of its energy, imparting it to the proton, as in the collision of two billiard balls. On the average it gives up more than half its energy at each collision, and so in about 20 collisions a 5 Mev neutron will be slowed down to thermal energies (0.1 electron volt or less). Thus a beam of fast neutrons after passing through about 6 to 10 cm of paraffin will be largely slowed down to thermal energies. At these low velocities neutrons are readily captured by a proton to form a deuteron, and so are quickly absorbed. A block of paraffin or a water cell is thus an effective barrier for neutrons—though they could easily pass through several feet of lead. The technique of handling neutron beams thus presents surprising problems, which there is not time to discuss.[4]

The problem of detecting and measuring neutron beams is also an interesting one. Since neutrons do not interact appreciably with atomic electrons they produce almost no ionization along their paths. Hence all the ordinary methods of detecting radiations (ionization chambers, counters, cloud chambers, etc.) which depend on ionization, are at first sight apparently useless. This difficulty is avoided by making use of collisions of neutrons with protons. The recoil protons produced in hydrogen-containing materials produce intense ionization, and the intensity of this can be taken as a measure of neutron intensity. The length of the recoil proton tracks produced in a hydrogen-filled cloud chamber can be used as a measure of the neutron velocities.

Still another method of detection is to make use of the disintegrations produced by neutrons in which alpha-particles are ejected, the alpha-particle ionization then being measured. The boron reaction is particularly useful:

$$B^{10}+n^1=Li^7+He^4.$$

This reaction is particularly efficient for low energy neutrons, for which the proton recoil ionization cannot be used. An ionization chamber lined with boron or filled with boron fluoride gas is widely used for the detection of neutrons.

Finally neutrons may be detected and measured by the amount of induced radioactivity they produce in certain materials such as silver.

It is evident from what has been said that the process by which a neutron loses energy in passing through matter is quite distinct from the processes by which gamma-rays or x-rays are absorbed. The latter give their energies to electrons while neutrons give their energy in appreciable quantities only to protons or other light nuclei. Electrons produce relatively small ionization over a long path; protons produce intense ionization over a short path; carbon nuclei, still more intense ionization over a still shorter path. Since it is to be expected that biological effects on individual cells will depend more on ionization *density* than total number of ions, it will be rather expected that neutrons will be more biologically effective than gamma-rays or x-rays of the same intensity. Preliminary experiments by Lawrence at California and by Zirkle of the Johnson Foundation indicate that this is indeed the case. Certainly this matter should be further investigated with great care and at once. As a matter of fact many experiments on neutron effects on both plant and animal tissue are now in progress, and there will be plenty of work in this field for many years to come.

There are difficulties in comparing neutron and x-ray effects which ought to be pointed out. X-ray doses are commonly measured in terms of the ionization produced in a standard chamber. Certain types of x-ray ionization chambers however may show almost no ionization for a neutron beam. A chamber used for neutrons must be lined with paraffin or Bakelite or other hydrogen-containing materials. The amount of ionization produced by a given beam will then depend greatly on the amount and arrangement of the materials introduced. Thus an x-ray and a neutron beam which produce equal ionization in one chamber will produce quite different effects in another one differently constructed. And the difference will depend on the energy (voltage) of the x-ray beam and the speed of the neutrons. On just what basis then can one say that neutrons are more effective than x-rays, or *vice versa*, unless there is some way of comparing intensities? Up to the present time each worker in this field has in general used a different type of ionization chamber, but in each case one has been chosen in which it was believed that the proton-recoil ionization would be comparable to what would be produced by the neutrons in biological materials. The results of different workers may be expected to be qualitatively comparable but it will soon be desirable to develop methods of quantitative measurement of neutron intensities relative to some arbitrary standard.

There is another important effect of neutrons which may be of biological interest. Both slow and fast neutrons are very effective in producing nuclear disintegrations, and in a large number of cases the disintegrations lead to products which are radioactive. (We shall have something to say about such materials later on.) In some cases also these disintegration reactions are accompanied by penetrating gamma-rays. A neutron is thus a "triple threat man" since in passing through biological materials it may simultaneously (1) produce recoil protons or other nuclei, (2) create radioactive atoms, (3) excite gamma-rays. It may be necessary to treat individually the biological results of these three processes. In most cases it seems likely that the first will be most important since disintegration probabilities are less than collision probabilities in general. And the elements for which the disintegration probability (giving rise to gamma-rays or radioactive products) is high are not materials commonly present in biological tissues, e.g., cadmium, samarium and boron. Nevertheless one should be on the lookout for cases in which neutrons produce effects on materials containing very little hydrogen, which may be due to disintegrations rather than collisions. In the case of very slow neutrons (energies of a few electron volts or

less) recoil ionization will be absent and their effects (if any) will be due solely to disintegrations.

It is possible that this property of neutrons of converting certain elements into their radioactive

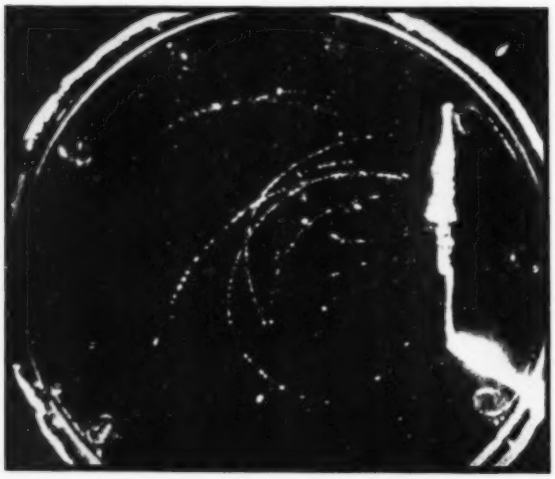


FIG. 2. Cloud chamber photograph showing the emission of positrons by radioactive nitrogen which was produced when the carbon plate shown at the right was bombarded by 900,000 electron-volt protons.

isotopes may be turned to good use. It might, for example, prove useful to direct a slow neutron beam at certain regions in the body and convert some of the elements present there into radioactive isotopes, whose radioactivity would remain for several hours after the neutron exposure had ceased. Since almost every element may be made radioactive by neutrons there are many possibilities to be considered. And this brings us directly then to a discussion of the possible uses in biological work of artificially produced radioactive materials.

Induced Radioactivity

It was in January 1934 that the Curie-Joliots reported that certain targets bombarded by alpha-particles continued to emit positrons for some time after the bombardment ceased. This positron activity was found to decay exponentially with time, just as in the case of the natural radioactive materials. By 1934 nuclear physics technique was quite well developed so that further studies of this induced radioactivity were undertaken at once by many laboratories. And now, a little over three years after the first announcement, over 220 radioactive isotopes are known, including practically every element of the periodic table and some beyond the end of the table. (Radioactive isotopes of elements of atomic number 93, 94, 95 and 96 are now known.)

The existence of these radioactive isotopes is easily understood. Each element in the periodic table has only a limited number (between 1 and 11) of stable isotopes which are found in nature. Thus fluorine has but one stable isotope, F^{19}; oxygen has three with weights, 16, 17, 18; and so on throughout the table, the heavier elements in general tending to have a larger number of stable isotopes. It is easy to see that many nuclear disintegration processes may give rise to isotopes of a particular element which are not found naturally and hence are presumably unstable. Thus one can produce F^{17}, F^{18} and F^{20}, all of which are unstable. Such atoms can in general convert themselves into stable ones by the ejection of a nuclear particle, and they will do so in the course of time. In most cases this is accomplished by the ejection of a positive or negative electron.[5] These electrons are expelled usually with considerable energy and the ionization they produce is readily measured. Each unstable atom has a certain probable lifetime, which is shorter the more unstable its nuclear structure. So, just as in the case of a large population of individuals, the number "dying" per unit time is proportional to the total number "living." And if none were being born the total population would decrease exponentially with time. The period or half-life is the time required for the number living (and also the death rate) to decrease to half its original value. This is just the law which governs the radioactivity of the natural radioelements, such as radium.

As an example if a C^{12} nucleus should capture a proton the resulting nucleus would be N^{13}, the charge and mass both being increased by 1. But N^{13} is not among the stable isotopes of nitrogen, though it could convert itself into the stable C^{13} by emitting a positron. This is what happens and it is found that N^{13} has a half-life of about 11 minutes. That is, if a large number of N^{13} atoms is formed at a given time the number of positrons emitted per second by the whole group falls to half its value in each succeeding 11 minute interval. It is in general possible by some nuclear

process to produce from one to six or more radioactive isotopes of every element in the periodic table (except hydrogen), and in many cases a particular isotope may be produced by several different nuclear reactions. As an example the radioactive Na^{24} with a period of 14.8 hours may be produced by bombarding the stable Na^{23} by either neutrons or deuterons, by bombarding Al by neutrons, or by bombarding Mg by neutrons or deuterons; five reactions in all. The properties of Na^{24} are the same no matter how it is produced. On decay the Na^{24} atom emits a negative electron and becomes stable Mg^{24}.

A few of the radioactive isotopes of some of the elements which occur commonly in biological materials are listed in Table I. The last column

TABLE I. *Some typical radioacitve isotopes.**

Atomic Number	Element and Mass	Emitted Particle	Stable Product	Half-life	Produced by
6	C^{11}	+	B^{11}	20.5 min.	B-*d*-*n*, B-*p*-γ
	C^{14}	−	N^{14}	? 3 mo.	N-*n*-*p*
7	N^{13}	+	C^{13}	11 min.	B-α-*n*, C-*d*-*p*, C-*p*-γ
8	O^{15}	+	N^{15}	126 sec.	N-*d*-*n*, N-*p*-γ
11	Na^{24}	−	Mg^{24}	14.8 hr.	Na-*d*-*n*, etc.
12	Mg^{27}	−	Al^{27}	10.2 min.	Mg-*d*-*n*, Al-*n*-*p*, —g-*n*-γ
15	P^{32}	−	S^{32}	14.5 days	P-*d*-*p*, S-*d*-α, P-*n*-γ, etc.
16	S^{35}	−	Cl^{35}	80 days	Cl-*n*-*p*
17	Cl^{38}	−	A^{38}	37 min.	Cl-*d*-*p*, Cl-*n*-γ, K-*n*-α
19	K^{42}	−	Ca^{42}	12.2 hr.	K-*n*-γ, K-*d*-*p*, etc.
20	Ca^{45}	−	Sc^{45}	2.4 hr.	Ca-*n*-γ, Ca-*d*-*p*, T-*n*-α
26	Fe?	?	Co	47 days	Fe-*d*-*p*
53	I^{128}	−	Xe^{128}	25 min.	I-*n*-γ
80	Hg^{205}	−	Tl^{205}	40 hr.	Hg-*n*-γ
82	Pb^{209}	−	Bi^{209}	8.5 days	Pb-*d*-*p*

* For a complete tabulation, see Livingston and Bethe, Rev. Mod. Phys. **9**, 359 (1937).

indicates how the isotopes are produced. For example the notation B-*d*-*n* opposite C^{11} indicates the bombardment of boron (B) by deuterons (*d*) with the ejection of a neutron (*n*), the reaction being

$$B^{10}+d^2 \rightarrow C^{11}+n^1.$$

The other notations are similar, the symbols *p*, *d*, *n*, α, γ standing respectively for proton, deuteron, neutron, alpha-particle (He nucleus) and gamma-ray. In general the bombarding particle must have an energy of from one to five million electron volts. To produce reasonable amounts of radioactive material intense beams of such particles must be available. This requires a cyclotron or some form of high voltage equipment.

The possibilities of using these materials in particular biological problems are so numerous that we can mention only a few by way of illustration. The more obvious possibilities (others are certain to come) may be grouped in two classes.

1. Therapeutic uses.—Radioactive isotopes of suitable elements may be injected into the body or directly into the organ to be treated. If it is found (and this requires much careful study) that the radiations they emit have definite therapeutic effects, then it is possible that induced radioactive isotopes may be more conveniently and effectively used than the naturally radioactive materials. Of particular interest is the possibility of selective irradiation of various organs of the body through the use of radioactive isotopes of elements which happen to be concentrated in those organs either under normal or diseased conditions. Thus one would use radioactive calcium or phosphorous for bone treatment, radio-iodine for the thyroid, etc., the active material going automatically to the spot where it is needed.

The clinical applications of these materials are thus both important and possibly spectacular. For this reason they have been perhaps prematurely hailed as a new boon to medicine. They certainly open up new possibilities, but their practical value is yet to be demonstrated and this will require a long period of careful research.

2. Tracing.—From the viewpoint of fundamental biology the use of radioactive materials as tracers is of far more interest than the therapeutic uses. In this field there are almost unbelievable possibilities which represent the most biologically promising of all aspects of nuclear physics. A prominent physiologist is quoted as stating his belief that the technique of using radioactive materials as tracers in biology may open a new era in that subject in as fundamental a way as did the invention of the microscope. The microscope made it possible to follow individual cells—the radioactive isotopes make it possible to follow individual atoms.

The uses to which this tool could be put are so many that it is as yet difficult even to classify them. Almost any chemical or biological problem in which it is desirable to follow the course of particular elements or compounds through a given system should be open to attack by this method. I will mention a few such problems

which happen to have come to my attention for purpose of illustration.

1. The role of phosphorus in bodily metabolism.[6] Phosphorus is an extremely important element to both plant and animals; phosphates appear in the teeth and bones, and the internal organs contain phosphorus in the form of organic compounds, such as the phospholipids. Many important questions arise, some of which can be answered by chemical methods and some not. How much of the phosphorus intake at a given time actually reaches the bones and teeth? How long after intake does it arrive at a given point? How rapid is the replacement? By what path through the bodily organs does it arrive? How are all of these things affected by abnormal conditions of diet or of disease? All of these questions and more can be answered by administering radiophosphorus—in the food or otherwise—and then following the radioactivity as it appears and disappears in various organs, the blood, the bones or teeth. Radiophosphorus (P^{32}) can be made in fairly strong samples by deuteron bombardment of ordinary phosphorus. The radiophosphorus can be combined into any compound after it is activated, and can therefore be fed or injected in the most convenient form. The half-life of P^{32} is about 14.8 days, so the activity of a given strong sample may be followed for many months. It decays with the emission of an electron, going to S^{32} which itself is probably harmless but in any case would be formed in amounts far too small to produce observable effects. By periodically testing the activity of samples taken from various parts of the body—either during the life of the animal or after killing and ashing—the amount and rate of phosphorus arrival may be determined with considerable precision. The many possibilities are sufficiently obvious as to require no further discussion before this group. Some work along this line has already been undertaken at various laboratories, with preliminary results which show clearly the power, accuracy and feasibility of the method. Some surprising results concerning phosphorus mobilization during certain diseases have already been indicated. I shall have to leave them to be discussed by those more familiar with biology than I am.

It is evident that similar studies could be made with almost any other element of interest; calcium, potassium, iodine, iron, etc. The only question which arises is whether there exists a radioactive isotope of the element in question whose period is sufficiently long to make the studies feasible. Here of course the physicist is somewhat helpless. He can make radioactive isotopes at will, but he has no control over the period of the resulting product; he must take what comes. Each isotope has its own characteristic period, determined by its own nuclear instability. The only hope is, if suitable periods of certain elements are not now known, that new isotopes will some day be discovered which are better adapted for such work. Nature has been fairly kind in this respect in providing so many periods of suitable length, and you may be sure that physicists will be busy for some time to come finding new ones.

2. Evidently a similar technique instead of being used on the body as a whole may be applied to studies of metabolism within individual organs, in nerves or muscles, even in individual cells. Many physiologists have already recognized these possibilities, and some have already planned programs along this line.

3. As the chemical nature of viruses, of hormones, and of various gland secretions become better known it should be possible to follow the behavior of these important but elusive agents, or to gain further information about their chemical composition.

4. There are many purely chemical or biochemical problems to which the tracing technique would seem applicable. Studies of chemical changes accompanying respiration would seem to be possible with the use of radio-oxygen—though the known periods unfortunately are rather short. The three-month period of carbon should, however, be extremely useful where CO_2 or organic compounds are involved. The possibility of tracing individual carbon atoms in various organic reactions might well open up a new field in organic and biochemistry.

I think these few examples will serve to illustrate the many possibilities—some of which will lie along lines which no one can now forsee.

From what has been said you will see that the

power of the radioactive tracing technique results from the following fortunate facts:

1. Radioactive isotopes of an element are chemically identical with the stable isotopes and will therefore behave in precisely the same way in all chemical or biological processes.

2. The active isotopes can be detected in extremely small quantities. In some cases it is possible to detect the presence in a given sample of as few as 1000 atoms of a given isotope. The presence of a million atoms would give accurately measureable results. This sensitivity results from the fact that impulse counters of various types, such as the Geiger-Müller tube counter, will register each individual β-ray (+ or − electron) entering it. If the geometrical conditions are properly chosen, from 0.1 to $\frac{1}{2}$ of all the β-rays arising in a given sample may be made to enter the counter. The fundamental law of radioactive decay is that the number of disintegrations occurring per second is proportional to the total number of atoms present, i.e.,

$$dN/dt = -\lambda N,$$

where $\lambda = 0.693/T$ where T is the half-life in seconds. Thus for $T = 1000$ sec. $= 17$ min. and for $N = 1000$

$$dN/dt = 0.693 \text{ disintegrations per sec.}$$
$$= 41.5 \text{ per minute.}$$

This is an observable number, even if only 1/10th of the β-rays produced enter the counter. In the case of P^{32} it is easy to prepare samples giving more than 10^6 disintegrations per minute or about 20,000 per sec. Since $T = 14.8$ days $= 1.3 \times 10^6$ sec., we find $N \leqq 3 \times 10^{10}$ radioactive atoms. This is only about 10^{-6} microgram. And if 0.01 percent of this amount appeared in any sample taken from an organism it could still be detected.

3. While it is possible that the presence of radioactive material in large quantities in certain organs might have definite biological effects on them due to the radiations emitted, and thus give abnormal results in metabolism experiments, it seems evident that in most cases the doses given may be made sufficiently small as to preclude any important effects due to the radiation itself.

4. Finally—and this should be strongly emphasized—the technique of detecting and measuring the radioactivity of any material is simple and capable of good precision. And when one recalls the days or weeks of work required to make a quantitative chemical analysis of substances containing a few micrograms of the element of interest, and the utter hopelessness of detecting changes involving only 10^{-12} to 10^{-14} grams, one sees the radioactive technique as child's play. And when one recalls that this relatively simple technique not only detects these small quantities of material, but actually distinguishes between the atoms already present (which were not radioactive) and those introduced at a particular time (which are) the possibilities become quite exciting. It is almost as though each individual atom carried a red flag to herald its presence. (It is even better than this since we know too little of the biological effects of red flags!)

The most common instrument which will likely be used in such investigations is the Geiger-Müller tube counter. This device, undoubtedly familiar to you, consists of a conducting cylinder along the axis of which is suspended a fine wire. A potential difference of the order of 1000 volts is applied between them. The whole is enclosed in a gas-tight envelope filled with air or other gas at a pressure of a few cm of Hg. A thin window must be provided for the entrance of slow β-particles. The counter wire is connected to a two- or three-stage amplifier which feeds a loud speaker or suitable registering equipment. The entry of a

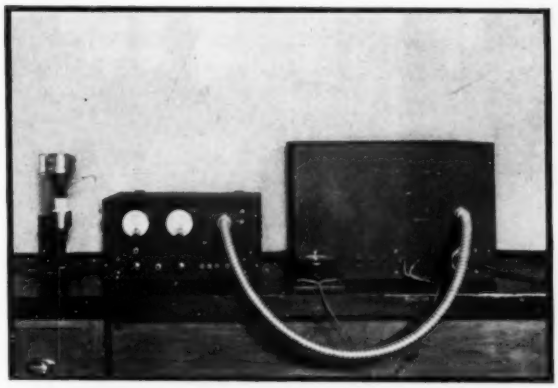


FIG. 3. The new biological "microscope" a Geiger-Müller tube counter used at the University of Rochester for biological work with radioactive materials.

β-ray into the cylinder causes a click in the speaker, or the pulse may be used to operate a mechanical register. Counter tube, rectifier, amplifier and registering device may all be contained in a box the size of a small radio set.[7] Observations consist merely in determining the number of registrations per unit time. Unfortunately such counters respond also to radioactive impurities which are always present in the air and in solid materials of the counter itself; and they also respond to cosmic rays. This means there will be a background of from 5 to 25 counts per minute which must be carefully determined and subtracted from the total observed with the source in place. It is this background which limits the ultimate sensitivity of the device, for the source being tested should produce a number of counts per minute at least as large as the background. Statistical fluctuations become important for low counting rates but can be reduced by extending the time of counting. A total of 10,000 counts must be made if the precision is to be as high as 1 percent, but only 100 counts will give 10 percent precision.

It is thus evident that simple readings of a clock and a registering device plus a little arithmetic is all that is involved in following atoms from place to place through even a complex biological system. It has been the experience that biologists who use this technique for the first time become wildly excited over the relatively simple way in which studies of most fundamental importance can be carried out. I shall expect this experience to be often repeated during coming years.

References

1. I am particularly indebted to Professor E. O. Lawrence, with whom these problems have been discussed in some detail.
2. The kinetic energy which a particle of charge e (electrostatic units) acquires in falling through a potential difference of V volts is given by the relation K.E. $= Ve/300$. An uncharged particle such as the neutron which has the same kinetic energy which an electron (or proton or any other single charged particle) would acquire in falling through a potential difference of 1 Mv is said to have an energy of 1 Mev.
3. *Inelastic* scattering (slowing down) of fast neutrons by heavy nuclei has recently been discovered at Cornell.
4. For a full account see Rasetti, *Nuclear Physics* (Prentice-Hall, 1936).
5. Electrons and positrons do not exist as such in the nucleus, but they may be created by the conversion of a proton into a neutron and a positron, or by the conversion of a neutron into a proton and an electron, the proper amount of energy being supplied or released.
6. The author is indebted to Dr. Wm. F. Bale of the University of Rochester School of Medicine for outlining the uses of radiophosphorus.
7. For further discussion and references see Harnwell and Livingood, *Experimental Atomic Physics* (McGraw-Hill, 1933) p. 412.

The cultural claims of science rest on the social fact that the use and misuse of science immediately affect the everyday life of every citizen in a modern community.—LANCELOT HOGBEN

Here and There

On February 17, 1938, Dr. Joseph Slepian was appointed to an associate directorship of the Westinghouse Research Laboratories in East Pittsburgh, Pennsylvania. Dr. Slepian's background has been that of an instructor at Cornell University, an apprentice in the motor division at Westinghouse, and later a consulting engineer in the research laboratory. Among his most important commercial contributions are a new lightning arrester and the deionizing circuit breaker. While he has become one of the country's outstanding authorities in the field of electrical engineering, he has also a great interest in higher mathematics, particularly in tensor analysis and integral equations.

★

Physics of the Universe

The University of Notre Dame announces the preliminary program of a Symposium on "The Physics of the Universe and the Nature of Primordial Particles" to be held Monday and Tuesday, May 2, and 3, 1938.

Public Lectures

The Basic Constituents of Matter. Carl D. Anderson, California Institute of Technology

Whence Cosmic Rays? Arthur H. Compton, University of Chicago

The Distribution of Matter in the Metagalaxy. Harlow Shapley, Harvard University

Technical Papers

Some Aspects of the Cosmic-Ray Problem. Carl D. Anderson, California Institute of Technology

The Nature of the Forces between Primordial Particles. Gregory Breit, University of Wisconsin

The Theory of Cosmic-Ray Particles. J. F. Carlson, Purdue University

Recent Research on Cosmic Rays. Arthur H. Compton, University of Chicago

The Relativistic Theory of Primordial Particles. Eugene Guth, University of Notre Dame

Cosmic Constants. Arthur E. Haas, University of Notre Dame

The Heat of the Stars and the Building of the Atoms in the Universe. William D. Harkins, University of Chicago

The Problem of the Expansion of the Universe. Canon Georges Lemaitre, University of Louvain, Visiting Professor at the University of Notre Dame

The Influence of the Magnetic Field of the Earth on Cosmic-Ray Particles. Manuel S. Vallarta, Massachusetts Institute of Technology

★

Acoustical Society of America

The Spring meeting of the Acoustical Society is to be held in Washington, D. C. on May 2–3, 1938. Address abstracts and all correspondence pertaining to the program to Mr. V. L. Chrisler, Department of Commerce, National Bureau of Standards, Washington, D. C. The present plans for the Washington meeting call for a symposium on acoustic standards. On Monday evening a trip is planned through the Department of Justice Building, where we will hear about the "G-Men" and their work. The Wardman Park Hotel will be headquarters.

★

For April

Polarized Light Enters the World of Everyday Life, by **M. GRABAU**

X-Rays in Agriculture, by **L. R. MAXWELL AND S. B. HENDRICKS**

Frederic Eugene Ives, by **J. A. TENNANT AND E. EPSTEAN**

Recent Progress in Supersonics, by **W. T. RICHARDS**

Order and Disorder in Alloys, by **H. A. BETHE**

Reaction Rate Theory of Viscosity and Some of Its Applications, by **R. H. EWELL**

Original research papers on the Reflection of Plane Waves and on the Elastic Parameters of Isotropic Solids.

Innovations in Instruments

Etching Solutions and Paste

Glass frosting materials for application to every kind of glassware have been developed by the Westinghouse Electric and Manufacturing Company. A glass etch solution has been formulated which will produce permanent white markings on glass. Being very fluid, it can be applied with a rubber stamp. It will dry rapidly with a small amount of heat. A metal etch solution for application by rubber stamp to brass, copper and silver has been developed. It produces a permanent and pleasing black finish, drying rapidly in the air without heat.

An acid glass frosting solution (No. 2) which is an etching material in paste form is now available. It produces a fine grained white satin frost on any type of glass in 10 to 15 seconds.

A safety frost solution for glass (No. 1) which has been developed is also in paste form. It produces a finish similar to that produced by No. 2, but requires treatment of the glass for 5 or 6 minutes, to produce similar results. Being less corrosive, fewer precautions need to be taken in its use. The relative slowness of its action is an operating advantage where the article is to be immersed in the solution, making it easy to obtain an even frost. Normally designs are applied to glass by coating the surface with paraffin or lacquer and cleaning off the portion which is to be etched.

★

New One-Ohm Standard Resistance

A new standard one-ohm resistance has been announced by the Leeds and Northrup Company of Philadelphia, Pa. This standard is said to have a high stability. It is adjusted in terms of the *absolute ohm* which will become effective as of January 1, 1940. In accordance with a preliminary announcement of the International Committee on Weights and Measures, this unit is taken as having the following relation to the present *international ohm:*

1 international ohm = 1.00048 absolute ohms.

Certificates may be obtained from the National Bureau of Standards, giving the value of a particular standard to one part in a million, in terms of the ohm as maintained by the Bureau.

★

Dynamic Balancer

A new machine has emerged from the research laboratory of the Westinghouse Electric and Manufacturing Company that allows the balancing of rotors "on location." The new balancer, which can be carried around in a couple of suitcases reduces the number of trial runs to two or three and simplifies the determination of correction weight position and amount to simple meter reading and elementary arithmetic. A sine wave alternator, little larger than an office fan motor, is connected to the shaft of the vibrating machine. The stator of this alternator can be shifted at will to control the phase relation. A vibration pickup is held against the vibrating body. The output of the pickup, suitably amplified, is fed to the voltage coil of an indicating wattmeter; the output of the sine wave alternator too goes to the current coil. By shifting the alternator stator to obtain zero and maximum wattmeter readings, the position and amount of correction weights for selected planes can be quickly prophesied.

★

Lightning and Thunder

Contrary to general belief, every lightning flash is not accompanied by a clap of thunder, according to K. B. McEachron, General Electric engineer in charge of high voltage and artificial lightning phenomena. His investigations have proved there are often lightning strokes which produce little or no thunder at all. Such flashes may appear just as bright as others, but their destructive force is less. Thunder is the result of a pressure wave caused by the sudden expansion of air created by a quick lightning discharge. All flashes do not release energy with the same speed. His studies during the past three years have revealed that in some cases, the electrical current is built up and compared to millionths of a second in other discharges. This so-called slow lightning produces no thunder. To the human eye it looks the same, and during a general storm the fact that one has heard no thunder in connection with any given stroke is generally overlooked.

★

New Source of Ultraviolet

A new, highly flexible source of ultraviolet radiation, consisting of a quartz-tube mercury vapor arc of improved efficiency mounted within a special reflector on a portable laboratory stand, has just been announced by the General Electric Vapor Lamp Company, Hoboken, N. J. The unit is self-starting and may be operated from any 120-volt or 220-volt a.c. line, through a portable auto-transformer supplied as part of the complete assembly. The sealed mercury vapor tube has high radiation characteristics in both the near and far ultraviolet zones, without the high heat losses and operating and maintenance problems inherent in open arcs. Rated operating life of the new tube is 2000 hours.

★

Arc Studies with High Speed Camera

Natural color motion pictures made with a high speed camera (1000 frames per second, with exposures of 1/10,000 second) of arcs being struck in different gases were shown

recently by Dr. C. G. Suits of the General Electric Research Laboratory. To the eye such arcs are simply explosive, spattering, intensely brilliant sparks. As slowed down and enlarged by the camera they were revealed as brilliantly colored flames, slowly pushing out tiny, brilliantly flashing globules of molten metal, and wandering gracefully around the edges of the electrodes for an interval before becoming more or less stable at the electrode tips. Flames of different colors were revealed by different gases and, in the case of oxygen, the fire was a conglomeration of a multitude of hues. The film also included views showing the arcs being blown sideways into graceful curves upon holding small but powerful Alnico magnets near them. Each of the sequences required more than a minute for showing; actually the action had been completed in less than a second.

Booklets Recently Received

Bulletin No. 835. Shallcross Mfg. Co., Collingdale, Pa., January, 1938. This bulletin contains descriptions of "Decade Resistance Boxes," "Decade Potentiometer Type 835," and "Standard Resistance Boxes." Bulletin No. 500 of the same company contains information about the "Shallcross Switches."

Castaloy. (7 pp.) Fisher Scientific Company, Pittsburgh, Pa. This issue contains information on the various uses of laboratory appliances.

Calibron Notebook. (24 pp.) No. 4. Published by Calibron Products, Incorporated, West Orange, New Jersey. This booklet contains a description of Metal Coating Processes now used, new developments, diffusion of metals and Fe-Al-Cl reaction.

General Radio Experimenter. (7 pp.) Vol. 12, No. 8, January, 1938. In this issue are articles on "Connection Errors in Capacitance Measurements," "A Visual Balance Indicator for A-C Bridge Measurements," and "Shielded Transformers for Balanced Circuit Measurements."

Oscillographer. (4 pp.) Vol. 1, Nos. 10–11, Dec. 1937–Jan. 1938. This issue contains a description of the "Phasmajector," a new source of video signals for use in demonstrating television, testing viewing cathode ray tubes or receiver performance.

Cenco News Chats. (14 pp.) Vol. III, No. 17, February, 1938. This issue contains articles on "A Question of Ethics," "Cenco Industrial Fellowships," "Photometer 'Goes to Town,'" "Lost Your Radium," "Colloidal Silica," "'Demonstration of the Year' Shown at Indianapolis," and "New High Pressure Mercury Arc."

Some of the New Books

Glossary of Physics, Compiled and Edited by Le Roy D. Weld. Pp. 255+x. McGraw-Hill Book Company, New York. Price $2.50.

Definitions, interpretations and exegetics of technical terms and discussions resulting therefrom never come to an end. Nor should they be expected to do so. As far as long used and established terms are concerned, it is important to acquaint the younger and more recent arrivals in a given field of the meaning of terms. Then too, because of improper instruction, or of legitimate differences in point of view, there is sure to be more or less looseness developing and "the lines of defense" must continually be repaired. Furthermore, in a living field of science new terms must continually be added to care for new concepts and ideas. It is only by trial and error, by frequent argumentation, or even by extended controversy, that agreement is finally reached on suitable terminology.

One can, then, in any given field welcome the appearance of a book which renders easier the refreshing of the mind or which frankly adds information to one's restricted range of knowledge. Such a book is the *Glossary of Physics*, compiled and edited by Le Roy D. Weld, professor of physics at Coe College. The Glossary had its origin in a collection of definitions gathered over a period of years by the editor. The collection was called to the attention of the Division of Physical Sciences of the National Research Council, and the compilation of the book was assisted by a grant-in-aid by the Council. Furthermore, the Council appointed the following Advisory Committee on a Glossary of Physical Terms:

A. W. Hull, General Electric Company
L. W. McKeehan, Yale University
F. K. Richtmyer, Cornell University
G. W. Stewart, University of Iowa
L. D. Weld, Coe College (*Chairman*)

The resulting book is chiefly the work of Dr. Weld who received the cooperation not only of the members of the Committee but of many consultants.

The Glossary contains over 3200 physical terms. Obviously this number cannot lay claim to being complete, but in the abridgment a careful choice has been made of the terms most likely to meet the need of a large number of readers. It is not limited to strictly physical terms but, in a useful manner, includes certain expressions or titles of certain phenomena such as "Davisson-Germer experiment," "Debye factor," etc. Also in the light of the extensive overlapping of various fields, such as physics and chemistry, many terms which are not strictly of the field of physics have been added.

In the making of a Glossary one may aim at conciseness or at definitiveness, or at both. In many cases these two aims are mutually exclusive, and one must decide which one is to be held in the foreground. Here the definitions

are designed to be clearly comprehensible and accurately concise, rather than laboriously detailed, and to be informative rather than authoritatively definitive. Undoubtedly in a manual of this scope that has been the wise course, for it is in obtaining agreement on definitive statements that the greatest difficulty arises. As a rule the concise statement is what the inquirer desires and, so far as this book is concerned, any lack in the other direction, so far as the meticulous inquirer is concerned, is largely taken care of by numerous references to more complete sources of information.

The Glossary should be an extremely useful one on the desk of any physicist and of any whose interest borders on and carries over into the field of physics.

P. I. WOLD
Union College, Schenectady, N. Y.

The Fundamental Principles of Quantum Mechanics—with Elementary Applications. EDWIN C. KEMBLE. Pp. 611+xviii, Figs. 27. McGraw-Hill, New York, 1937. Price $6.00.

Professor Kemble has given us in his *Fundamental Principles* a book which stands out far above the average run of "elementary" texts on quantum mechanics, and which may well be expected to take a place along with the accepted classics of the subject. The book is characterized by a conscientious and praiseworthy effort to present a decent, if not entirely rigorous, mathematical formulation of the physical problems and of their solutions. In the author's own words, he has "been led to try [his] hand at bridging the gap between the exacting technique of von Neumann and the usual less rigorous formulations of the theory;" in the opinion of the reviewer, he has achieved a large measure of success in this arduous undertaking.

Chapters I–III, V contain an historical account of the physical origins of the theory, and a treatment of the standard elementary problems, including that of the hydrogenic atom. Chapters IV, VI–VIII, which occupy almost one-third of the book, are devoted to the general mathematical theory of operators with discrete and continuous spectra, with application to dynamical variables of greatest physical importance—coordinate, linear and angular momentum, energy and symmetry operators. These chapters are followed by chapters on each of the topics: theory of measurement, matrix theory, perturbation theory and the more elementary aspects of radiation theory. The concluding Chapters XIII, XIV constitute an introduction to the general theory of the structure and spectra of many-electron atoms. Little space is given to collision problems, and the relativistic theory of the electron and quantum electrodynamics are omitted as beyond the scope of the book.

In the opinion of the reviewer, the excellent features of the book are not particularly enhanced by the emphasis which the author places throughout on the "operational point of view." No exception can be taken to this viewpoint insofar as it refers to the more considered scientific methodology which has been forced upon physics by its extension into new domains, as in the relativity and quantum theories; as such, it offers an appropriate, philosophically neutral, setting for the facts and theories of physics. But, at a time when philosophers and physicists alike are tying themselves into hard knots over the philosophical implications of modern physics, the physicist, *qua* physicist, should take care not to elevate this "point of view" to the status of a positivistic dogma, to the exclusion of other possible philosophical interpretations. And Kemble has, in effect, injected such issues into his account of quantum mechanics—as in his discussions of measurement, in which he says (p. 328) "the wave function is merely a subjective computational tool and not *in any sense* a description of objective reality" (italics ours)! Whether because of this attitude, or because of his attempt to treat the observing mechanism on a classical basis, Kemble's discussion of the important question of the "reduction of the wave packet" (sec. 41d) seems incomplete and even unsatisfactory; the reader would be well advised to supplement this section with the excellent accounts of measurement in Pauli's *Handbuch* article and in von Neumann's monograph.

The book as a whole is heartily to be recommended to the serious student of quantum mechanics; the chapters on the more rigorous formulation of the theory should prove especially valuable to anyone who is seeking a deeper understanding of these fundamental points, which are too often glossed over for the sake of the rapid acquirement of manipulative facility.

H. P. ROBERTSON
Princeton University

Light, Principles and Experiments. GEORGE S. MONK. Pp. 477+xi, Figs. 269. McGraw-Hill Book Company, Inc. New York and London, 1937. Price $5.00.

This book is intended as an intermediate text for students who have completed an ordinary college course in general physics and a course in differential and integral calculus. It includes Geometrical Optics (8 chapters, 99 pages); Physical Optics (9 chapters, 240 pages); Laboratory Directions for 23 experiments (74 pages); Problems, at the end of nearly every chapter; eight Appendices; six Tables of Useful Data, with logarithms and trigonometric functions; and Answers to Problems. The fruit of long experience and of tremendous labor, it constitutes a welcome and much needed addition to available textbooks at intermediate level. The inclusion of geometrical optics, physical optics, and laboratory instructions, all in one volume, will make it especially valuable.

After a period of five years in which there have been no new intermediate textbooks in optics, there now appear two in the same year, both by the same publisher. The book under review is one; the other is Jenkins and White's *Fundamentals of Physical Optics*, which, as the title implies, limits itself to physical optics.

The illustrations in Monk's *Light* are ample in number and excellent in quality; they have evidently been the object of great care. There are few errors in the figures and also few in the text, either of typography or of statement. The treatment is comprehensive and carries its conclusions far. The book should prove extremely good, for students who come well prepared and for whom a thorough course

is in prospect—a course of at least four hours per week for a semester. Under other circumstances, difficulty may arise. For a shorter course or for students who are only moderately well prepared, the treatment may prove too advanced and, in fairly numerous places, less lucid than it should be. But it is no doubt too much to expect that a book should be thorough and go far, and yet be always simple and clear. The author is to be congratulated upon the extent to which he has achieved the former of these desiderata.

CHARLES F. MEYER
University of Michigan

Contributors to This Issue

After obtaining a B.S. in physics from Case School of Applied Science in 1933, L. P. Tarasov pursued graduate work in the physics and metallurgy departments at M.I.T. and received a D.Sc. in Applied Physics in 1937. Since last fall, he has been on the staff of the Research Laboratory of the General Electric Company doing work in ferromagnetism.

Gabriel Kron was born in Hungary in 1901. He graduated from the University of Michigan in 1924. Later he worked with various electrical manufacturing companies as a machinist, draughtsman, designer, research-, development-, and consulting engineer. He is now with the General Electric Company, Schenectady. He has written many articles on electrical engineering subjects.

Miss Adelaide Easley received an A.B. degree from Colorado College in 1924 and an M.A. degree from Northwestern University in 1926. She left Northwestern University in 1928 to join the Nela Park Research Laboratory of the General Electric Company, and since 1929 has been employed as physicist in the Lamp Development Laboratory of the General Electric Company, Cleveland, Ohio.

Donald D. Hinman was born in Ravenna, Ohio in 1909. He received an A.B. degree from Oberlin College in 1931. Since 1932 he has been employed as physicist in the Lamp Development Laboratory of the General Electric Company, Cleveland, Ohio.

Dr. W. E. Forsythe received his B.S. degree from Dennison University in 1907 and his Ph.D. from the University of Wisconsin in 1911. In 1913 he joined the staff of the Incandescent Lamp Department of the General Electric Company. Since 1929 his title has been that of physicist in the Lamp Development Laboratory, Nela Park, Cleveland, Ohio. He is the author of many papers on high temperature measurements and radiation.

Contributed Original Research

Quantitative Measurements of Texture by the Magnetic Torque Method *

L. P. Tarasov†
Massachusetts Institute of Technology, Cambridge, Massachusetts

Magnetic torque measurements offer a possibility of specifying texture in ferromagnetic materials in a quantitative fashion and several methods have been proposed for accomplishing this. The method of Akulov and Brüchatov, although formally correct, is shown to be so restricted by its assumptions as to be inapplicable even to the commonest texture. A method proposed by Bitter, involving root-mean-square deviations from some mean orientation, is found to fail in practice because of the uncertainty in the measurement of the magnetic energy. A modification of this method, in which x-ray data are used to establish a relationship between two of the root-mean-square deviations, gives results in fair agreement with the pole figure of a cold-rolled iron-silicon alloy.

THE pole figure method of representing preferred orientation, such as is observed in cold-rolled and in recrystallized metals, gives the experienced reader a fairly good picture of the distribution of the particle orientations with respect to some reference plane. This method was invented by Wever[1] over ten years ago and is now finding considerable use.[2] It is as yet only qualitative, since present-day x-ray technique is not sufficiently well developed to warrant more than a few degrees of shading in the pole figure, corresponding to the ranges of intensity observed in the x-ray film. In spite of the inability of the experimenter to furnish us with a group of numbers descriptive of the texture, his qualitative map or pole figure is satisfactory for most purposes. When the texture is fairly simple, it is often found convenient, in explaining the texture to those not acquainted with pole figures, to talk in terms of some mean orientation with deviations therefrom of a certain number of degrees in each of several directions. These deviations, even though expressed in degrees, are obviously no more quantitative than the pole figure, for the information given by these deviations is only a part of that given by the pole figure.

It is highly desirable to be able to express textures quantitatively, because slight variations in texture can be described more concisely and accurately this way than by drawing several pole figures. Bozorth[3] has done this, using x-rays, for deviations occurring in the texture of cold-rolled iron. The fact that ferromagnetic disks give magnetic torque curves similar to those for single crystals whenever there is some preferred orientation, has led Akulov and Brüchatov[4] and Bitter[5] to propose methods of utilizing the numerical values associated with these torque curves in specifying the texture.

The Method of Akulov and Brüchatov

Akulov and Brüchatov have attempted to specify texture in terms of so many percent of each of two ideal orientations, the rest being made up of randomly oriented material. The choice of the two from the four possible orientations given below is limited by the nature of the

* Part of a doctoral thesis submitted to the physics department of M.I.T. in October 1937.
† Now at the Research Laboratory, General Electric Co., Schenectady, N. Y.
[1] F. Wever, Trans. A. I. M. E. **93**, 51 (1931).
[2] C. S. Barrett, Trans. A. I. M. E. **124**, 29 (1937).
[3] R. M. Bozorth, Phys. Rev. **50**, 1076 (1936).
[4] N. Akulov and N. Brüchatov, Ann. d. Physik **15**, 741 (1932).
[5] F. Bitter, *Introduction to Ferromagnetism* (McGraw-Hill, 1937) pp. 213–222.

torque curve. The four orientations are:

In rolling plane	{100}	{100}	{110}	{110}
Along rolling direction	<100>	<110>	<100>	<110>

Unfortunately, this method can be applied to the torque curve for nearly any sample so as to give a definite numerical result, but in many cases the answer obtained will be wrong according to x-ray data.

As an example, consider a typical cold-rolling texture for an iron-silicon alloy for which the coefficients A_1 and A_2 of the torque curve can be taken as -0.05 and $+0.40$, respectively.[6] These coefficients are defined by the equation

$$T/a_F = A_1 \sin 2\theta + A_2 \sin 4\theta, \qquad (1)$$

where T is the torque measured in ergs/cm^3, a_F is the anisotropy constant[5] in the same units and θ is the angle in the plane of the disk measured from the rolling direction. According to the equations of Akulov and Brüchatov and the values of the coefficients just quoted, the texture is composed of 47½ percent of the second orientation, 10 percent of the third, and the rest random. However, the absence of (110) planes in the rolling plane is a well established fact for cold-rolled iron and the 10 percent of the third type of orientation cannot be present. The restriction of the possible orientations to those listed is necessary for mathematical reasons, but the rejection of all intermediate orientations which always occur in practice to a greater or lesser extent makes this method useless when attempts are made to apply it. It is the intermediate orientations which contribute to A_1, while the method outlined above assumes that there must be (110) planes in the rolling plane whenever A_1 is not zero. The intermediate orientations also make their contributions to A_2 with the result that any answer obtained in this way has little real significance.

Root-Mean-Square Deviations from Ideal Orientation

A more reasonable approach to a quantitative study of preferred orientation is to assume an average ideal orientation as determined from a pole figure and then to use the torque curve as a means of finding the r.m.s. (root-mean-square) deviations from the ideal orientation. The theory on which this method rests was developed by

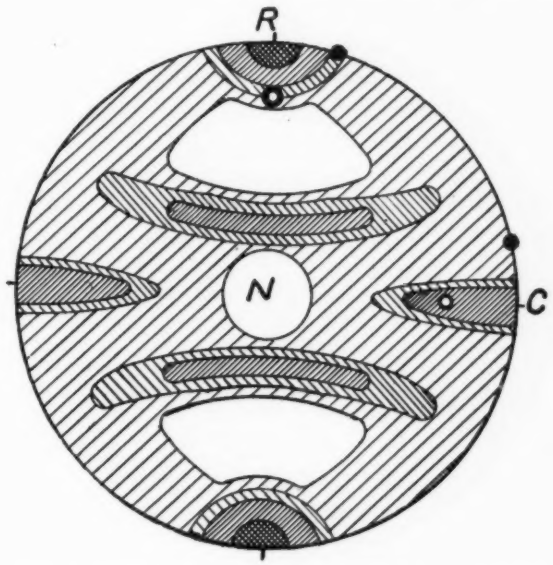


FIG. 1. {110} pole figure for cold-rolled iron containing 2 percent silicon (adapted from paper by Barrett, Ansel and Mehl, reference 9). The root-mean-square deviations for the sample discussed in the text, and calculated on the assumption that $c=n$, are shown as follows: open circle, $r=18.9°$; filled circle, $c=14.3°$; solid circle, $n=14.3°$.

Bitter[5] and the work described below was done in order to determine the possibilities and limitations of the method.

The basic idea is that if the mean orientation is known, the magnetic energy E_θ can be expressed in terms of the r.m.s. deviations around each of three mutually perpendicular axes. These consist of the rolling direction R, the cross direction C (at right angles to R in the rolling plane), and N, the normal to the rolling plane. If ω_R, ω_C and ω_N are the angular rotations in radians of the ideal orientation around these axes, and if we set $[\omega_R^2]_{Av}=r^2$ for short, then for the mean orientation of the cold-rolled texture which has (001) normal to N and [110] parallel to R,

$$E_\theta/a_F = \tfrac{1}{4}(1-c^2-r^2-8n^2)\cos 4\theta + \tfrac{1}{2}(c^2-r^2)\cos 2\theta + \tfrac{1}{4}(1+3c^2+3r^2) \qquad (2)$$

and

$$T/a_F = (1-c^2-r^2-8n^2)\sin 4\theta + (c^2-r^2)\sin 2\theta.$$

[6] L. P. Tarasov, reported at the New York Meeting, A.I.M.E., February, 1938.

The derivation is strictly valid only for ω's small enough that the errors in the approximations $\sin\omega=\omega$ and $\cos\omega=1-\frac{1}{2}\omega^2$ can be neglected. Since there are three unknowns and a torque curve can give only two independent equations in the unknowns, one for A_1 and the other for A_2, the energy equation has to be used to provide a solution. The only convenient way of using E_θ is to make a magnetization curve with the magnetic field in some definite direction such as R. In this case E_θ is the area on the graph of I against H between the observed curve and the one for the $<100>$ direction of a single crystal of the same material.

Provided the magnetization is along R, the energy equation can be simplified to read $E_\theta/a_F=\frac{1}{2}+c^2-2n^2$. A further simplification is to let $E_\theta/a_F-\frac{1}{2}=U$, $1-A_2=V$ and $A_1=W$. In this new notation,

$$\begin{aligned} c^2-2n^2&=U\\ c^2+r^2+8n^2&=V\\ c^2-r^2&=W \end{aligned} \qquad (3)$$

and the solution of these equations is easily found to be

$$\begin{aligned} 6c^2&=4U+V+W\\ 6r^2&=4U+V-5W\\ 12n^2&=-2U+V+W \end{aligned} \qquad (4)$$

By using the extreme values of U, V and W, it can be shown that the maximum values of the mean square deviations consistent with Eq. (4) are $1/8$ for n^2, $5/18$ for c^2, and $25/36$ for r^2, corresponding to r.m.s. values of about 20°, 30°, and 40°, respectively. The upper limits to the r.m.s. deviations set by the conditions of the derivation of the original equation for E_θ are even lower than those given, but if a liberal error is permissible, then textures can be investigated for which the r.m.s. deviations are not greater than those listed.

Equation (4) shows directly the relative sensitiveness of the mean square deviations to changes in E_θ and the harmonic coefficients. Thus c^2 and r^2 are both very sensitive to a change in E_θ/a_F and only one-fourth as sensitive to a similar sized change in A_2; also n^2 is much less sensitive than the other two deviations, especially with respect to E_θ/a_F.

Application of Bitter's Method

A strip of 3 percent silicon-iron, 12″ long and 1″ wide, that had been cold-rolled 80 percent down to 20 mils (described elsewhere[6]) was annealed for one hour at 480°C in air so as to remove as much of the strain as possible without causing any recrystallization. This is always a necessary part of the procedure, because the energy as given in Eq. (2) refers to completely strain-free material. It is impossible to relieve all the strains without recrystallizing the material, and this residual strain introduces an error into the magnetization curve which can be at least approximately corrected when the curve is sheared so as to bring the initial straight portion vertical. This is the accepted method for correcting experimental observations on *strain-free* single crystals for the effect of the demagnetizing field. The introduction of strain into the problem results in this straight line portion of the curve making a larger angle with the I axis than it would otherwise, so that shearing the curve as described should take care, not only of the demagnetizing field, but also of most of the residual strain. The effect of strain upon the curve beyond the first break from the vertical is unknown, but should be quite small since the strain has itself been made small by the annealing treatment.

The intensity of magnetization was measured directly with a relative accuracy of one part in several thousand, in fields as high as 950 oersteds, saturation being attained at about two-thirds of this. The area between the corrected magnetization curve and the straight lines $H=0$ and $I=I_s$ was found to be 7.0×10^4 ergs/cm^3, I_s being the saturation value of I. From this area there has to be subtracted E_{100}, the energy of magnetization in the $<100>$ direction. As this was unknown for the chemical composition of the strip, a value for iron was sought in the literature, but the agreement between various experimenters was very poor. For example, Honda, Masumoto and Kaya[7] found (as derived by Piety[8]) that for a single crystal of iron of just medium purity $E_{100}=0.60\times10^4$ ergs/cm^3, whereas Piety,[8] using

[7] K. Honda, H. Masumoto and S. Kaya, Tôhoku Univ. Sci. Reports **17**, 111 (1928).
[8] R. G. Piety, Phys. Rev. **50**, 1173 (1936).

two carefully annealed single crystals of much greater purity found values of 2.48×10^4 and 1.33×10^4. It should be noted that Piety's two crystals were grown from the same material and under the same circumstances.

Since the values are so scattered for one composition, we shall arbitrarily use the entirely reasonable values of 1.0×10^4 and 1.5×10^4 ergs/cm³ and see what effect an error in E_{100} can be expected to cause. Letting $a_F=1.67\times10^5$ ergs/cm³ (this was the only accurately known factor), a value of -0.141 was found for U, assuming $E_{100}=1.0\times10^4$. The torque curve gave $V=0.673$ and $W=-0.046$. Substitution into Eq. (4) gave these values of the mean square deviations: $c^2=0.010$, $r^2=0.056$, $n^2=0.076$. These correspond to r.m.s. values expressed in degrees of $c=5.7°$, $r=13.6°$, and $n=15.8°$. Changing E_{100} from 1.0×10^4 to 1.5×10^4 ergs/cm³ caused c^2 to become negative. The values calculated according to the new assumption were $c^2=-0.010$, $r^2=0.036$, and $n^2=0.081$; in other words c became imaginary, r became 10.9° and n changed but slightly to 16.3°.

This method is thus seen to give answers which depend more upon a totally unknown factor than upon the texture of the specimen. Even the first answer does not correspond to the texture given by the pole figure, for it seems that according to the latter c and n are approximately the same, while r should be several times as large. This is evident from Fig. 1, in which is shown a pole figure for the {110} planes adapted from the one given by Barrett, Ansel and Mehl[9] for 95 percent cold-rolled iron containing 2 percent silicon. The reason for the disagreement probably arises from the presence of a considerable quantity of random orientation which not only affects the r.m.s. deviations as it properly should, but also changes the course of the magnetization curve so as to bring it nearer the one for perfectly random orientation. The conclusion that must inevitably be drawn is that this method has no practical usefulness until the problem of E_{100} is thoroughly investigated; after that is done, it may be possible to straighten out the other difficulties.

[9] C. S. Barrett, G. Ansel and R. F. Mehl, Trans. A. I. M. E. **125**, 516 (1937).

The question may arise as to why we could not measure the magnetization curve in two different directions and find the area between them so as to make a knowledge of E_{100} unnecessary. An inspection of Eq. (2) shows that taking the difference between any two energies results in the cancellation of the constant term, and it is this term which gives us our third independent measurable quantity necessary for the solution of an equation in three unknowns.

Modified Quantitative Method

Since all the trouble about applying the quantitative method has arisen from the errors involved in the correct determination of the energy, it is interesting to modify the method so as to make use only of the torque data assisted by x-ray data. The pole figures generally show that as a first approximation the rotations around c and n can be set equal. In such a case, $c^2=n^2$ and the last two equations in Eq. (4) become $9c^2+r^2=V$ and $c^2-r^2=W$. The solution of these is

$$\begin{aligned} 10c^2&=V+W \\ 10r^2&=V-9W \end{aligned} \qquad (5)$$

Substitution of the values of V and W previously stated makes $c=14.3°$ and $r=18.9°$. This answer would have required that $E_\theta=7.3\times10^4$ ergs/cm³ on the basis of the theory previously developed.

The new answers look fairly reasonable as shown in Fig. 1, where these values of c, r, and n are indicated (in one quadrant only) by the three types of circles. Although c and n are somewhat too large, this can be attributed to the presence of a considerable amount of randomly oriented texture. The large value of n can be accounted for in this manner, but c is too large simply because we made the arbitrary assumption that $c=n$. Finally, the circle for r is seen to fall in just about the expected position.

The assumption that $c=n$ was arbitrary, but seemed entirely justified on the basis of the heavily shaded regions of the pole figure. Since the lightly shaded regions of the pole figure also seem to be of importance, we can modify our previous assumption by letting c be some fraction of n. For example, we can let $2c=n$. In this case we find that $c=7.8°$, $n=15.6°$, and $r=14.5°$

The change in the ratio of c to n causes c to drop to half its former value and r to decrease a little, while n remains the same. The agreement with the pole figure is now definitely better.

The possibility of making use of this modified method depends mostly upon whether or not any meaning can be attached to some definite ratio of c to n. If such a quantity does exist for a series of samples, then variations in c could be detected. It will be necessary to do some careful x-ray work on a more or less quantitative basis before the existence of any such ratio can be verified.

ACKNOWLEDGMENT

The author wishes to express his thanks to Mr. A. R. Kaufman, of the department of metallurgy, M.I.T., for the use of his apparatus and for his help in measuring the magnetization curve.

Invariant Form of the Maxwell-Lorentz Field Equations for Accelerated Systems

GABRIEL KRON
Engineering General Department, General Electric Company, Schenectady, N. Y.
(Received October 8, 1937)

It is shown that the invariant field equations of Maxwell for stationary bodies are also valid for accelerated bodies, if all ordinary differentiations are replaced by absolute differentiation. The affine connection $\Gamma^{\gamma}_{\alpha\beta}$ appearing in the absolute derivatives refers to a general non-Riemannian metric space, with a holonomic or nonholonomic reference frame. The field equations are in a more general form than those used in the five-dimensional unified field theory and reduce to the latter as a special case. However, the presence of an electrostatic field is not assumed in this paper. The motion of an accelerated body possessing mass, electric current and magnetic flux is given by an extended form of the dynamical equations of Lagrange containing the same general affine connection as the field equations. The material terms of the dynamical equations are expressed by a mechanical impulse-energy tensor and the resultant motion of the material body under the influence of an electromagnetic field and a mechanical force is expressed by a combined impulse-energy equation. The equations are calculated for a representative rotating electrical machine used in industry.

INTRODUCTION

IN a previous article[1] it was shown that the performance of a quasi-holonomic dynamical system, in which both mechanical and electromagnetic energies coexist, may be represented by the motion of a point in an n-dimensional non-Riemannian, metric space. It was also shown that from geometrical considerations a more general form of Lagrangean equation of motion can be derived, that reduces as a special case to the usual Lagrangean form for holonomic axes and to the Boltzman-Hamel form for nonholonomic axes.

By using the affine connection defined by the *extended* form of the Lagrangean equation, in these pages the usual tensor forms of the field equations of Maxwell are generalized. The following cases may be considered, that also occur in the various types of rotating electrical machinery used in industry:

(1) When the reference frame is *rigidly* connected to the conductors (moving or stationary) the affine connection is

$$\Gamma^{\gamma}_{\alpha\beta} \overset{h}{=} \tfrac{1}{2}a^{\gamma\delta}\left(\frac{\partial a_{\alpha\delta}}{\partial x^{\beta}}+\frac{\partial a_{\delta\beta}}{\partial x^{\alpha}}-\frac{\partial a_{\alpha\beta}}{\partial x^{\delta}}\right) \overset{h}{=} \left\{ \begin{matrix} \gamma \\ \alpha\beta \end{matrix} \right\}, \quad \text{(I)}$$

being equivalent to the holonomic Christoffel symbol.

[1] Kron, "Quasi-Holonomic Dynamical Systems," Physics 7, 143 (1936). In the present paper also the notation and terminology of Schouten and Struik's book, *Einführung in die neueren Methoden der Differentialgeometrie* (1935) is used.

(2) When the reference frame has a *constant uniform velocity* with respect to the conductors, the affine connection is

$$\Gamma^{\gamma}{}_{\alpha\beta}=\tfrac{1}{2}a^{\gamma\delta}\left(\frac{\partial a_{\alpha\delta}}{\partial x^{\beta}}+\frac{\partial a_{\delta\beta}}{\partial x^{\alpha}}-\frac{\partial a_{\alpha\beta}}{\partial x^{\delta}}\right)+a_{\alpha\delta}a^{\gamma\epsilon}\Omega^{\delta}{}_{\beta\epsilon}$$

$$+a_{\beta\delta}a^{\gamma\epsilon}\Omega^{\delta}{}_{\alpha\epsilon}-\Omega^{\gamma}{}_{\alpha\beta}=\left\{\begin{matrix}\gamma\\ \alpha\beta\end{matrix}\right\}, \quad \text{(II)}$$

where $\Omega^{\gamma}{}_{\alpha\beta}$ is a nontensor geometric object, the "nonholonomic object." The whole expression is equivalent to the nonholonomic Christoffel symbol $\left\{\begin{matrix}\gamma\\ \alpha\beta\end{matrix}\right\}$.

(3) When there is a relative *accelerated motion* between the reference frame and the conductors, the affine connection is

$$\Gamma^{\gamma}{}_{\alpha\beta}\overset{h}{=}\left\{\begin{matrix}\gamma\\ \alpha\beta\end{matrix}\right\}+S_{\alpha\beta}{}^{\gamma}+S_{\beta}{}^{\gamma}{}_{\alpha}-S^{\gamma}{}_{\alpha\beta}, \quad \text{(III)}$$

where $S_{\alpha\beta}{}^{\gamma}$ is the skew-symmetric "torsion tensor" and $\left\{\begin{matrix}\gamma\\ \alpha\beta\end{matrix}\right\}$ the holonomic Christoffel symbol defined in Eq. (I).

(4) When there is a relative *accelerated motion* between the reference frame and the conductors on one hand and a relative *constant, uniform velocity* between the various axes, or between the various conductors, the affine connection is

$$\Gamma^{\gamma}{}_{\alpha\beta}=\left\{\begin{matrix}\gamma\\ \alpha\beta\end{matrix}\right\}+S_{\alpha\beta}{}^{\gamma}+S_{\beta}{}^{\gamma}{}_{\alpha}-S^{\gamma}{}_{\alpha\beta}, \quad \text{(IV)}$$

where $\left\{\begin{matrix}\gamma\\ \alpha\beta\end{matrix}\right\}$ is the nonholonomic Christoffel symbol defined in Eq. (II).

In setting up the field equations time will be assumed as an additional variable, $t=x^{t}$. Free charges ρ and the electrostatic field D^{τ} will not be considered since they do not play an important part in the performance calculation of rotating electrical machinery. Their role will be considered in another publication. Rotating magnetic media do not introduce anything new and in general only case (3) with the components of the Christoffel symbol all zero will be considered in greater detail in the paper.

It is found that *the introduction of a relative acceleration between the reference frame and the conductors is equivalent to the appearance of a new magnetic flux-density wave B_{τ} in space, the rotor flux-density wave, that is nonexistent in the absence of an accelerated reference frame* (as far as measurements from the reference frame are concerned). The presence of this new magnetic field is manifested by the appearance of a new tensor, the "torsion tensor" $S_{\alpha\beta}{}^{\gamma}$, which is also nonexistent in the absence of an accelerated reference frame.

A relative *constant, uniform velocity* between the reference frame and the conductors however does not produce a new tensor, only a new nontensor geometric object, the "nonholonomic object" $\Omega^{\gamma}{}_{\alpha\beta}$.

The question of the existence or nonexistence of the rotor flux-density wave B_{τ} has led to two schools of thought in the performance analysis of electrical machinery, one group using flux-density waves in space, the other using flux-linkages (vector potential φ_{α}) measured at the machine terminals.

It will be assumed that the reader is acquainted with the first five sections of the previous article. The same notation and terminology is used here.

The following indices will be employed:

(1) *electrical axes:*
  (a) variable: $m, n, k, \cdots$
  (b) fixed: *d and q*
(2) *geometrical axes:*
  (a) variable: $u, v, w, \cdots$
  (b) fixed: s
(3) *electrical and geometrical axes:*
  (a) variable: $\pi, \sigma, \rho, \cdots$
  (b) fixed: d, q and s
(4) *electrical, geometrical and time axes:*
  (a) variable: $\alpha, \beta, \gamma, \cdots$
  (b) fixed: d, q, s, and t

In a general electromechanical system any number of electrical and geometrical axes may occur. In the paper an example with the minimum number of axes is calculated.

I. The Equation of Motion of Lagrange

It has been shown in a previous article that for quasiholonomic dynamical systems the *equation of motion* of Lagrange may be written for holonomic or nonholonomic reference frames as

$$\frac{d^2x^{\tau}}{dt^2}+\Gamma^{\tau}{}_{\sigma\tau}\frac{dx^{\sigma}}{dt}\frac{dx^{\tau}}{dt}+R_{\sigma}{}^{\tau}\frac{dx^{\sigma}}{dt}=f^{\tau}. \quad (1.1)$$

The affine connection is defined[2] in Eqs. (16) and (17) as

$$\Gamma^{\tau}{}_{\sigma\tau}=\left\{\begin{matrix}\pi\\ \sigma\tau\end{matrix}\right\}+S_{\sigma\tau}{}^{\tau}+S_{\tau}{}^{\tau}{}_{\sigma}-S^{\tau}{}_{\sigma\tau}, \qquad (1.2)$$

where $\left\{\begin{matrix}\pi\\ \sigma\tau\end{matrix}\right\}$ is the holonomic or nonholonomic form of the Christoffel symbol and $S_{\sigma\tau}{}^{\tau}$ is the torsion tensor, skew-symmetric in its first two indices. In terms of the latter the equation of motion is

$$\frac{d^2x^{\tau}}{dt^2}+\left\{\begin{matrix}\pi\\ \sigma\tau\end{matrix}\right\}\frac{dx^{\sigma}}{dt}\frac{dx^{\tau}}{dt}-2S^{\tau}{}_{\sigma\tau}\frac{dx^{\sigma}}{dt}\frac{dx^{\tau}}{dt}+R^{\tau}{}_{\sigma}\frac{dx^{\sigma}}{dt}=f^{\tau}. \qquad (1.3)$$

In order to establish a physical picture, the equation of motion should be expressed separately along the electrical and the geometrical axes, the free indices being covariant indices.

Assuming the *electrical* indices m, n, k as free indices, the equation of voltage, representing the motion of electric charges, is (neglecting frictional forces)

$$\underbrace{a_{mn}\frac{d^2x^n}{dt^2}+[un,\,m]\frac{dx^u}{dt}\frac{dx^n}{dt}}_{\text{induced voltage}}$$

$$\underbrace{-2S_{mun}\frac{dx^u}{dt}\frac{dx^n}{dt}}_{\text{generated voltage}}=\underbrace{f_m}_{\text{impressed voltage}}. \qquad (1.4)$$

The induced voltage term includes the voltages due to the:

(1) variation of currents,
(2) motion of reference axes,
(3) motion of magnetic paths.

The generated voltage is due to the motion of the conductors.

Assuming the *geometrical* indices u, v as free indices, the equation of torque (representing the motion of the rotor and salient poles) is

$$\underbrace{a_{uv}\frac{d^2x^v}{dt^2}}_{\text{inertial torque}}+\underbrace{[mn,\,u]\frac{dx^m}{dt}\frac{dx^n}{dt}}_{\text{torque on salient pole}}$$

$$\underbrace{-2S_{umn}\frac{dx^m}{dt}\frac{dx^n}{dt}}_{\text{torque on rotor}}=\underbrace{f_u}_{\text{impressed torque}}. \qquad (1.5)$$

II. The Equations of Power

If each of the above equations is multiplied by $dx^{\tau}/dt=\dot{x}^{\tau}$, the equations of power are:

(a) for the whole system

$$\underbrace{a_{\tau\sigma}\ddot{x}^{\sigma}\dot{x}^{\tau}+[\sigma\tau,\,\pi]\dot{x}^{\sigma}\dot{x}^{\tau}\dot{x}^{\pi}}_{\text{time rate of change of total stored energy}}=\underbrace{f_{\tau}\dot{x}^{\tau}}_{\text{power input}}. \qquad (2.1)$$

The power flow due to the presence of the torsion tensor is zero.

(b) for the electrical axes

$$\underbrace{a_{mn}\ddot{x}^n\dot{x}^m+[un,\,m]\dot{x}^u\dot{x}^n\dot{x}^m}_{\text{time rate of change of stored electrical energy}}$$

$$\underbrace{-2S_{mun}\dot{x}^m\dot{x}^u\dot{x}^n}_{\text{electrical output}}=\underbrace{f_m\dot{x}^m}_{\text{electrical input}}. \qquad (2.2)$$

(c) for the geometrical axes

$$\underbrace{a_{uv}\ddot{x}^v\dot{x}^u+[mn,\,u]\dot{x}^m\dot{x}^n\dot{x}^u}_{\text{time rate of change of stored mechanical energy}}$$

$$\underbrace{-2S_{umn}\dot{x}^m\dot{x}^n\dot{x}^u}_{\text{mechanical output}}=\underbrace{f_u\dot{x}^u}_{\text{mechanical input}}. \qquad (2.3)$$

III. Assumptions

It will be assumed that no free electric charges ρ exist in the rotating machine, nor permanent magnets, nor electrostatic flux-density D^{τ}. Since an electric field intensity E_{τ} is impressed on the machine, *time will be introduced as an extra variable*, represented by $t=x^4$. It also follows that $dx^4/dt=dt/dt=1$. When time is one of the fixed indices, the variable indices are $\alpha, \beta, \gamma\cdots$.

It will be assumed that if a cross section is made of any rotating electrical machine, at any

[2] Banesh Hoffmann, "Projective Relativity and the Einstein-Mayer Unified Field Theory," Phys. Rev. **43**, 615 (1933), Eq. (13).

one instant the current-density, flux-density, impressed voltage, etc., waves are distributed sinusoidally in space in each concentric layer of winding of the rotor (shown schematically in Fig. 1). The concentric windings are interconnected with stationary or moving junction-points in various manner.

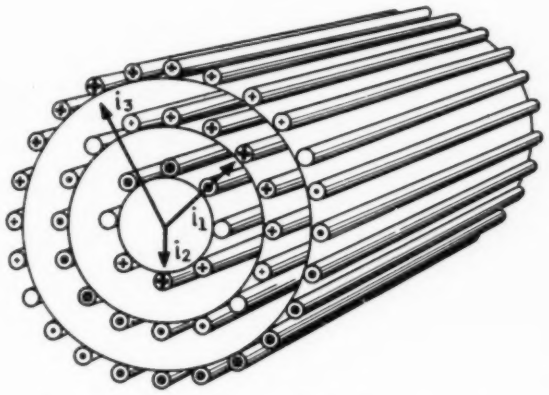


FIG. 1. Concentric layers of windings on a rotor.

Any rotating electrical machine may be assumed to consist of a collection of 0-cells, 1-cells and 2-cells having accelerated motion with respect to each other. In order to avoid carrying along matrices with large number of rows and columns, the invariant equations developed will be illustrated for a representative machine having a single 2-cell (a single rotor winding). Its design constants are assumed to be calculated with two 1-cells (two stator windings) permanently short-circuited, as shown in Fig. 2. Numerous other examples for a large variety of industrial machines are worked out in another publication.[3]

The rotor will have two stationary axes d and q at right angles to each other in space. A space quantity like B_π will represent a sinusoidal space wave having projections B_d and B_q along the d and the q axes.

The following form of the torsion tensor as defined in the previous article (with time axis added) will often be needed.

$$-2S_{\alpha s}{}^{\beta} = -2\Omega_{s\alpha}{}^{\beta} = A_\alpha{}^k A_s{}^s \partial A_k{}^\beta / \partial x^s =$$

α \ β	d	q	s	t
d	0	1	0	0
q	−1	0	0	0
s	0	0	0	0
t	0	0	0	0

$= 2S_{s\alpha}{}^{\beta} =$ (3.1)

It should be noted that one of the first three indices of the torsion tensor is the geometrical index s, the others are electrical indices d and q. If two of the fixed electrical indices are the same, the components are zero.

The starting point will be the field equations of Maxwell valid for *stationary* bodies, containing ordinary derivatives (or for moving bodies containing absolute derivatives with the holonomic Christoffel symbol $[\alpha\beta, \gamma]$ as the affine connection, which however drops out of the tensor equations). The invariant equations for accelerated motion between the reference frame and the conductors will be established by *replacing each ordinary derivative by an absolute derivative*, using the affine connection $\Gamma_{\alpha\beta,\,\gamma}$ developed for the dynamical equation of Lagrange in Eq. (1.2).

Absolute derivatives with respect to x^t will be taken as if x^t were t, that is first the absolute differential, say $\delta\varphi_a$, is taken, then both sides

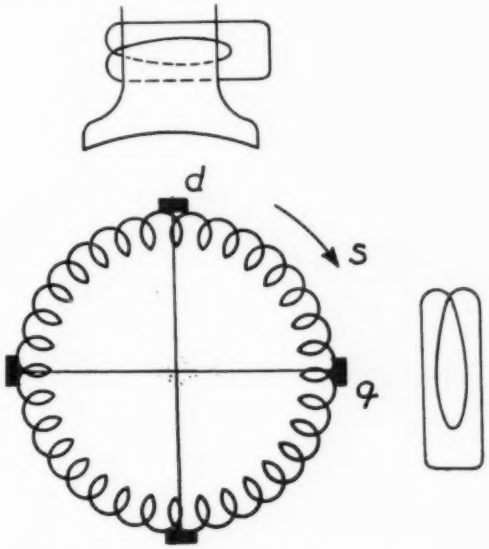


FIG. 2. Quasi-holonomic reference axes.

[3] Kron, "The Application of Tensors to the Analysis of Rotating Electrical Machinery," Gen. Elec. Rev. April, 1935—December, 1937.

are divided by dt. For instance

$$\delta\varphi_\alpha = d\varphi_\alpha - \Gamma^\beta{}_{\alpha\gamma}\varphi_\beta dx^\gamma, \tag{3.2}$$

$$\frac{\delta\varphi_\alpha}{\partial x^t} = \frac{\delta\varphi_\alpha}{dt} = \frac{d\varphi_\alpha}{dt} - \Gamma^\beta{}_{\alpha\gamma}\varphi_\beta \frac{dx^\gamma}{dt}. \tag{3.3}$$

Hence it is possible to define

$$S_{t\alpha}{}^\beta = S_{s\alpha}{}^\beta \partial x^s/\partial x^t, \tag{3.4}$$

$$2S_{t\alpha}{}^\beta = \begin{array}{c|cccc} \alpha \backslash \beta & d & q & s & t \\ \hline d & 0 & \partial x^s/\partial x^t & 0 & 0 \\ q & -\partial x^s/\partial x^t & 0 & 0 & 0 \\ s & 0 & 0 & 0 & 0 \\ t & 0 & 0 & 0 & 0 \end{array} \tag{3.5}$$

analogously to $2S_{s\alpha}{}^\beta$, so that

$$\delta\varphi_\alpha/\partial x^t = \partial\varphi_\alpha/\partial x^t - \Gamma^\beta{}_{\alpha t}\varphi_\beta. \tag{3.6}$$

The usual Maxwell's equations in the symbolism of vector analysis will be represented here with the indices $\pi, \sigma, \epsilon \cdots$ corresponding to the fixed indices d, q, s. That is two types of vectors will have to be differentiated:

(1) Those that occur in the usual Maxwell's equation when time t is not an extra variable. These vectors will have the indices $\pi, \sigma, \epsilon \cdots$.

(2) Those that occur in the tensor formulation with time as an additional variable x^t. These vectors will have the indices $\alpha, \beta, \gamma \cdots$.

Also the following vectors will be used

$$dx^m/dt = i^m; \quad dx^u/dt = v^u; \quad dx^t/dt = 1, \tag{3.7}$$

where i^m represents currents, v^u represents velocities of the rotor, the magnetic paths, or the reference axes.

IV. The Vector Potential

Maxwell defines the "*vector potential*" (a three-dimensional vector) A_π, as the *resultant* flux linking each circuit

$$A_\pi = a_{\pi\sigma} i^\sigma. \tag{4.1}$$

When the stator windings are assumed to be short circuited and the magnetic paths stationary, the metric tensor is defined as

$$a_{\pi\alpha} = \begin{array}{c|ccc} \pi \backslash \alpha & d & q & s \\ \hline d & L'_d & 0 & 0 \\ q & 0 & L'_q & 0 \\ s & 0 & 0 & 1 \end{array} \tag{4.2}$$

where L'_d and L'_q are the short circuit inductances of the rotor windings. In case of moving salient poles the components are functions of x^s.

Hence the invariant "vector potential" is defined as

$$\varphi_\alpha = \begin{array}{c|cccc} \alpha & d & q & s & t \\ \hline & \varphi_d & \varphi_q & 0 & 0 \end{array} \tag{4.3}$$

$$\varphi_\alpha = \begin{array}{c|cccc} \alpha & d & q & s & t \\ \hline & L'_d i^d & L'_q i^q & 0 & 0 \end{array}$$

Each component of φ_α can be measured at the machine terminals.

When free charges ρ exist, the time component of φ_α contains the scalar potential V. Its component along s is zero.

V. The Field Tensor $F_{\alpha\beta}$

If the vector potential φ_α is known, the field tensor is found in stationary bodies (or in moving bodies with holonomic reference axes) by

$$F_{\alpha\beta} = \partial\varphi_\alpha/\partial x^\beta - \partial\varphi_\beta/\partial x^\alpha. \tag{5.1}$$

Hence replacing ordinary derivatives by absolute derivatives

$$F_{\alpha\beta} = \delta\varphi_\alpha/\partial x^\beta - \delta\varphi_\beta/\partial x^\alpha = 2\nabla_{[\beta}\varphi_{\alpha]}. \tag{5.2}$$

$$F_{\alpha\beta} = \partial\varphi_\alpha/\partial x^\beta - \partial\varphi_\beta/\partial x^\alpha - 2S_{\alpha\beta}{}^\gamma\varphi_\gamma. \tag{5.3}$$

If the reference frame is also nonholonomic, Eq. (5.3) becomes

$$F_{\alpha\beta} = \partial\varphi_\alpha/\partial x^\beta - \partial\varphi_\beta/\partial x^\alpha - 2S_{\alpha\beta}{}^\gamma\varphi_\gamma + 2\Omega_{\alpha\beta}{}^\gamma\varphi_\gamma, \tag{5.4}$$

where $\Omega^\gamma{}_{\alpha\beta}$ is the nonholonomic object.

The expression is skew-symmetric in the indices α and β. The right-hand side does not become zero in the following two cases:

(1) $\beta = s$ (geometrical axis)

$$F_{as} = -2S_{as}{}^{\delta}\varphi_{\gamma}. \tag{5.5}$$

Multiplying the vector (4.3) with matrix (3.5)

$F_{as}=$	α: d	q	s	t
	φ_q	$-\varphi_d$	0	0

(5.6)

representing a vector at right angles in space to the flux-linkage vector. It will be called the *flux-density vector* $-B_s$ so that

$F_{as}=-B_s=$	π: d	q	s
	$-B_d$	$-B_q$	0

(5.7)

In the presence of *moving magnetic paths* when the metric tensor a_{as} is a function of x^s

$$F_{as} = \partial\varphi_a / \partial x^s - 2S_{as}{}^{\gamma}\varphi_{\gamma} + 2\Omega_{as}\varphi_{\gamma}, \tag{5.8}$$

$-B_s=$	π: d	q	s
	$\partial\varphi_d/\partial x^s + \varphi'_q + \varphi''_q$	$\partial\varphi_q/\partial x^s - \varphi'_d - \varphi''_d$	0

F_{as} has a nonzero component along its geometrical axis s in the presence of free charges ρ.

Equation (5.5) can be written as

$$B_s = \text{Abs. curl } \varphi_s, \tag{5.9}$$

(2) $\beta = t$ (time axis),

$$F_{at} = \delta\varphi_a / \partial x^t = \partial\varphi_a / \partial x^t - 2S_{at}{}^{\gamma}\varphi_{\gamma} + 2\Omega_{at}\varphi_{\gamma}. \tag{5.10}$$

The right-hand side of this equation is identical with the equation of voltage of Lagrange, Eq. (1.4) if φ_{γ} is replaced by $a_{\gamma m}i^m$. Hence

$$F_{at} = -E_s = \frac{\partial\varphi_s}{\partial t} - 2S_{s\gamma}{}^{\sigma}\frac{dx^{\gamma}}{dt}\frac{dx^{\sigma}}{dt} + 2\Omega_{s\sigma}\frac{dx^{\sigma}}{dt}\frac{dx^{\sigma}}{dt}. \tag{5.11}$$

$F_{at}=$	α: d	q	t	s
	e_d	e_q	0	0

(5.12)

It corresponds to the Maxwell-Lorentz equation for moving bodies

$$\mathbf{E} = -d\mathbf{A}/dt + \mathbf{B}_1 \times \mathbf{v}_1 + \mathbf{B}_2 \times \mathbf{v}_2, \tag{5.13}$$

where $\mathbf{v}_1$ = variable velocity and $\mathbf{v}_2$ = constant velocity.

Hence the field tensor $F_{a\beta}$ as defined by Eq. (5.3) for the representative machine is

$F_{a\beta}=$	α \ β: d	q	s	t
	d	0	0	$-B_d$	$-E_d$
	q	0	0	$-B_q$	$-E_q$
	s	B_d	B_q	0	0
	t	E_d	E_q	0	0

(5.14)

$F_{a\beta}=$	α \ β: d	q	s	t
	d	0	0	φ_q	e_d
	q	0	0	$-\varphi_d$	e_q
	s	$-\varphi_q$	φ_d	0	0
	t	$-e_d$	$-e_q$	0	0

(5.15)

$F_{a\beta}=$	α \ β: d	q	s	t
	d				
	q			$-B_\pi$	$-E_\pi$
	s		B_π		
	t		E_π		

(5.16)

VI. The First Field Equation

Since $F_{a\beta}$ is a skew-symmetric tensor and is the absolute curl of φ_a, it automatically follows that

$$\nabla_{[\alpha}F_{\beta\gamma]}=\nabla_\alpha F_{\beta\gamma}+\nabla_\beta F_{\gamma\alpha}+\nabla_\gamma F_{\alpha\beta}=0 \quad (6.1)$$

or
$$\delta F_{\beta\gamma}/\partial x^\alpha+\delta F_{\gamma\alpha}/\partial x^\beta+\delta F_{\alpha\beta}/\partial x^\gamma=0. \quad (6.2)$$

Let $\alpha=t$ (time axis), $\gamma=s$ (geometrical) and $\beta=$electrical axis (either d or q)

$$\delta F_{ms}/\partial x^t=\delta F_{mt}/\partial x^s-\delta F_{st}/\partial x^m$$
$$=\delta E_s/\partial x^m-\delta E_m/\partial x^s. \quad (6.3)$$

By Eq. (5.4)

$$\delta F_{ms}/\partial x^t=-\delta B_\pi/dt=\partial E_s/\partial x^m-\partial E_m/\partial x^s$$
$$-2S_{sm}{}^kE_k+2\Omega^k{}_{sm}E_k. \quad (6.4)$$

This is equivalent to

$$\text{Abs. curl } E_m=-\frac{\delta B_m}{dt}=-\frac{dB_m}{dt}$$
$$+2S_{ms}{}^kB_k\frac{dx^s}{dt}-2\Omega^k{}_{ms}B_k\frac{dx^s}{dt}, \quad (6.5)$$

$$\text{Abs. curl } \mathbf{E}=-d\mathbf{B}/dt-\text{Abs. curl } \mathbf{B}\times\mathbf{v}, \quad (6.6)$$

representing the generalization of the following Maxwell-Lorentz equation for moving bodies

$$\text{Curl } \mathbf{E}=-d\mathbf{B}/dt-\text{curl } \mathbf{B}\times\mathbf{v}. \quad (6.7)$$

Although the last equation is in the form given by Hertz, it is equivalent to the slightly different form given by Lorentz.

Equation (6.5) gives the same scalar equations (along the missing index) as the equation of voltage (5.10), except it expresses the voltages in terms of flux-densities B_m instead of flux-linkages φ_m. This is not surprising since Eq. (6.5) is a consequence of Eq. (5.10).

Let none of the indices assume the time axis and let $k=s$. Then

$$\delta F_{ns}/\partial x^m+\delta F_{sm}/\partial x^n=0. \quad (6.8)$$

To find out what this represents, let $n=d$ and $m=q$

$$\delta F_{ds}/\partial x^q+\delta F_{sq}/\partial x^d=0. \quad (6.9)$$

By Eq. (5.5) it represents

$$\delta\varphi_q/\partial x^q+\delta\varphi_d/\partial x^d=\delta\varphi_m/\partial x^m=\partial\varphi_m/\partial x^m=0. \quad (6.10)$$

That is Eq. (6.8) represents

$$\text{Abs. Div. } \mathbf{A}=0. \quad (6.11)$$

VII. The Field Tensor $H^{\alpha\beta}$

In ponderable bodies the relation between the magnetic field-intensity vector H^π and the magnetic flux-density vector B_π is

$$B_\pi=\mu_{\pi\sigma}H^\sigma \quad \text{or} \quad \mathbf{B}=\boldsymbol{\mu}\cdot\mathbf{H}. \quad (7.1)$$

The permeability tensor is, if the salient poles are stationary

π \ σ	d	q	s
d	L'_q	0	0
$\mu_{\pi\sigma}=q$	0	L'_d	0
s	0	0	1

(7.2)

so that by Eqs. (5.6) and (4.1)

$$B_\pi=$$

π	d	q	s
	$-i^qL'_q$	$i^dL'_d$	0

(7.3)

$$\mathbf{H}=H^\sigma=\mathfrak{R}^{\sigma\pi}B_\pi=$$

π	d	q	s
	$-i^q$	i^d	0

(7.4)

where $\mathfrak{R}^{\sigma\pi}$ is the inverse of $\mu_{\pi\sigma}$. The field-intensity vector H^σ is identical with the *magneto-motive-force* wave, at right angles in space to the current-density wave.

The relation between B_π and H^π may also be written as

$$B_\pi=H_\pi+M_\pi \quad \mathbf{B}=\mathbf{H}+\mathbf{M} \quad (7.5)$$

where M_π is the magnetic polarization vector (induced, not permanent)

$$M_\pi=B_\pi-H_\pi=$$

π	d	q	S
	$i^q(L'_d-L'_q)$	$-i^d(L'_d-L'_q)$	0

(7.6)

and it represents the effect of saliency around the periphery of the machine. *The difference between H^π and H_π should be noted.*

Similar relations exist between E_π and D^π, but in these pages D^π is assumed to be zero.

Analogously to the field tensor $F_{\alpha\beta}$ (which is built up from B_π and E_π) another skew-symmetric field tensor $H^{\alpha\beta}$ can be built up from H^π and D^π for the representative machine as

$H^{\alpha\beta}=$

$\alpha \backslash \beta$	d	q	s	t
d	0	0	$-H^d$	$-D^d$
q	0	0	$-H^q$	$-D^q$
s	H^d	H^q	0	0
t	D^d	D^q	0	0

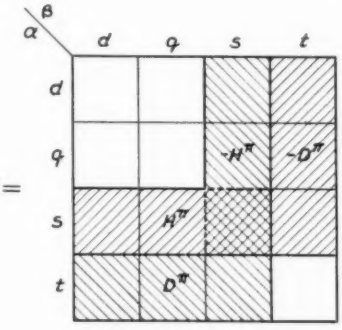
(7.7)

$H^{\alpha\beta}=$

$\alpha \backslash \beta$	d	q	s	t
d	0	0	i^q	0
q	0	0	$-i^d$	0
s	$-i^q$	i^d	0	0
t	0	0	0	0

(7.8)

$H^{\alpha\beta}=$



(7.9)

VIII. The Second Field Equation

The second independent field equation is

$$\delta H^{\alpha\beta}/\partial x^\beta = s^\alpha. \tag{8.1}$$

The right-hand side follows from the following consideration

$$\delta H^{\alpha\beta}/\partial\alpha^\beta = \partial H^{\alpha\beta}/\partial\alpha^\beta + S_{\gamma\delta}{}^\alpha H^{\gamma\delta}. \tag{8.2}$$

By Eqs. (3.4) and (3.5)

$$S_{\gamma\delta}{}^\alpha H^{\gamma\delta} = S_{s\delta}{}^\alpha H^{s\delta} + S_{\gamma s}{}^\alpha H^{\gamma s} \tag{8.3}$$

$$= -2S_{s\gamma}{}^\alpha H^{\gamma s} = s^\alpha, \tag{8.4}$$

since calculation gives, by multiplying Eq. (7.8) by Eq. (3.1)

$-2S_{s\gamma}{}^\alpha H^{\gamma s}=$

α	d	q	s	t
	i^d	i^q	0	0

$= s^\alpha$, (8.5)

representing the *rotor currents* (conduction currents).

Hence the generalized field equation is

$$\text{Abs. Curl } \mathbf{H} = \mathbf{i}. \tag{8.6}$$

In the absence of D^π and ρ no other field equation is involved in (8.1)

The *equation of continuity* is

$$\delta s^\alpha/\partial x^\alpha = \partial s^\alpha/\partial x^\alpha + \Gamma^\alpha{}_{\beta\alpha} s^\beta = 0. \tag{8.7}$$

In the absence of electric charges ρ, it is equivalent to

$$\text{Abs. Div. } \mathbf{i} = 0. \tag{8.8}$$

IX. The Force-Power Equation

The invariant equation of Maxwell

$$F_{\alpha\beta} s^\beta = p_\alpha \tag{9.1}$$

consists of two parts:

(1) Let the free index $\alpha = s$ (geometrical axis)

(9.2)

$$F_{s\beta} s^\beta = p_s,$$

$$p_s = B_d i^d + B_q i^q = B_\pi i^\pi. \tag{9.3}$$

It represents the negative of the *electromagnetic torque* developed by the machine, given usually in the form

$$\mathbf{p} = \mathbf{B} \times \mathbf{i}. \tag{9.4}$$

(2) Let the free index $\alpha = t$ (time axis)

$$F_{t\beta} s^\beta = p_t, \tag{9.5}$$

$$p_t = E_d i^d + E_q i^q = E_\alpha i^\alpha. \tag{9.6}$$

It represents the negative of the *electromagnetic power input* into the machine, usually written as

$$\mathbf{p} = \mathbf{E} \cdot \mathbf{i}. \tag{9.7}$$

X. The Stress-Energy Tensor

If in the equation of force (9.1) s^α is replaced by its value in Eq. (8.1) a new tensor is defined, the "stress-energy tensor" (or "impulse-energy tensor")

$$T_\alpha^{\ \beta} = F_{\alpha\beta} H^{\beta\gamma} - \tfrac{1}{4}\delta_\alpha^{\ \beta} F_{\gamma\delta} H^{\gamma\delta}. \qquad (10.1)$$

If the values of $F_{\alpha\beta}$ and $H^{\alpha\beta}$ given in Eqs. (5.14) and (7.7) are substituted, the stress-energy tensor of the representative machine is

$T_\alpha^{\ \beta} =$

α \ β	d	q	s	t
d	$B_d H^d - W$	$B_d H^q$	0	0
q	$B_q H^d$	$B_q H^q - W$	0	0
s	0	0	W	0
t	0	0	$E_d H^d + E_q H^q$	$-W$

(10.2)

The instantaneous *stored magnetic energy* is

$$T_t^{\ t} = \mathbf{W} = -\tfrac{1}{2}(B_d H^d + B_q H^q)$$
$$= -\tfrac{1}{2} B_\pi H^\pi = -W. \qquad (10.3)$$

This term is not a function of the geometrical variable x^s.

The terms along the axes d, q and s are the *Maxwell stresses*

$$T_\pi^{\ \sigma} = \boldsymbol{\pi} = B_\pi H^\sigma - \tfrac{1}{2}\delta_\pi^{\ \sigma} B_\epsilon H^\epsilon. \qquad (10.4)$$

The remaining term is the "*Poynting's vector*"

$$T_t^{\ s} = \boldsymbol{\Xi} = e_d i^q - e_q i^d = E_\pi H^\pi, \qquad (10.5)$$

representing the magnetic power flowing into the machine.

It should be noted that the momentum tensor

$$T_\pi^{\ t} = \mathbf{G} = B_\pi D^\pi \qquad (10.6)$$

is zero, since $D^\pi = 0$.

Hence the stress-energy tensor can also be written as

$T_\alpha^{\ \beta} =$

α \ β	d	q	s	t
d				
q		$\boldsymbol{\pi}$		$\mathbf{G}$
s				
t		$\boldsymbol{\Xi}$		$\mathbf{W}$

(10.7)

If the time axis t is not considered, the remaining tensor must represent solely the Maxwell stresses $\boldsymbol{\pi}$.

The following important relation is satisfied

$$T_\alpha^{\ \alpha} = 0. \qquad (10.8)$$

The part representing the Maxwell stresses is asymmetrical.[4] Its skew-symmetric part is

$T_\pi^{\ \sigma} - T^\sigma_{\ \pi} =$

π \ σ	d	q	s
d	0	$-p$	0
q	p	0	0
s	0	0	0

(10.9)

representing the torque developed by the machine

$$-p = i^q \varphi_d - i^d \varphi_q = B_d H^q - B_q H^d. \qquad (10.10)$$

This term exists because the B_π and H_π vectors are not in space phase. Also

$$T_\pi^{\ \pi} = 0. \qquad (10.11)$$

It should be noted that to get a physical picture the mixed form $T_\alpha^{\ \beta}$ has to be considered instead of $T_{\alpha\beta}$ or $T^{\alpha\beta}$.

XI. The Electromagnetic Impulse-Energy Equation

The equation of force and power (9.1), becomes in terms of the stress-energy tensor

$$\delta T_\alpha^{\ \beta} / \partial x^\beta = p_\alpha. \qquad (11.1)$$

[4] Dupont, "La Force et le Couple electro-magnetiques dans un champ gravifique," Bull. de l'Academie Royale de Belgique (Classe des Sciences), (1936) No. 2.

(1) Let the free index $\alpha = s$ (geometrical axis)

$$\delta T_s^{\beta}/\partial x^{\beta} = -S_{s\beta}{}^{\gamma} T_{\gamma}^{\beta} - S_{s\gamma}{}^{\beta} T^{\gamma}{}_{\beta} = p_s = -p. \quad (11.2)$$

If the calculation is performed, p_s is found to be the negative of the torque developed by the machine

$$\text{Abs. Div. } \overset{m}{\pi} = -\mathbf{p}. \quad (11.3)$$

(2) Let the free index $\alpha = t$ (time axis)

$$\delta T_t^{\beta}/\partial x^{\beta} = \partial T_t^{\beta}/\partial x^{\beta} - \Gamma^{\gamma}{}_{t\beta} T_{\gamma}^{\beta}. \quad (11.4)$$

When the dummy index $\beta = s$ (geometrical)

$$\delta T_t^{s}/\partial x^{s} = -S_{t\beta}{}^{\gamma} T_{\gamma}^{\beta} - S_{t\gamma}{}^{\beta} T^{\gamma}{}_{\beta} = p_s v^s = -pv \quad (11.5)$$

representing that part of the electrical power input that changes into mechanical power, that is

$$\text{Abs. Div. } \overset{m}{\mathfrak{S}} = -\mathbf{p}v. \quad (11.6)$$

Assuming the dummy index $\beta = t$,

$$\delta T_t^{t}/\partial x^{t} = \delta \mathbf{W}/\partial x^{t} = -\partial W/\partial t, \quad (11.7)$$

representing the negative time rate of change of stored magnetic energy.

The two components (11.6) and (11.7) may be combined into one equation of power

$$\delta \mathbf{W}/dt + \text{Abs. Div. } \overset{m}{\mathfrak{S}} = -\mathbf{E} \cdot \mathbf{i}. \quad (11.8)$$

Hence p_α in Eq. (11.1) represents the negative electrical torque and the negative electrical power input.

XII. The Mechanical Impulse-Energy Equation

The Maxwell-Lorentz equations do not incorporate the forces and power due to the presence of *material bodies*, that is, the moment of inertia of the rotor and the applied mechanical force do not appear in the field equations. Considering the dynamical equations of Lagrange, Eqs. (1.1)–(2.3), the field equations of Maxwell-Lorentz give:

(1) The equation of voltage (1.4),

(2) The torque exerted by the magnetic field, that is the torque on the conductors and on the magnetic material, given in Eq. (1.5),

(3) The electrical equation of power (2.2).

However, the field equations do not include:

(1) The remaining part of the equation of torque (1.5),

(2) The mechanical equation of power (2.3).

It is possible to set up for the purely mechanical system certain field equations that are analogous to the purely electrical field equations. Let the following tensors be defined:

(1) The mechanical momentum is

$$\overset{m}{\mathbf{G}} = mv = a_{ss} v^s. \quad (12.1)$$

(2) The stored *mechanical energy* is

$$\overset{m}{\mathbf{W}} = -\tfrac{1}{2} mv^2 = -\tfrac{1}{2} a_{ss} v^s v^s, \quad (12.2)$$

being a function of time only, but not of x^s.

(3) The flux of energy-flow, the "*Poynting's vector*" is, for a body moving with a velocity v,

$$\overset{m}{\mathfrak{S}} = Sv = Sv^s, \quad (12.3)$$

where S is the kinetic energy stored in the rotor due to the flow of the electrical energy from the electromagnetic field.

$$S = \int p_s dx^s. \quad (12.4)$$

(4) The stress tensor $\overset{m}{\pi}$ is

$\overset{m}{T}_{\pi}{}^{\sigma} = \overset{m}{\pi}$

π \ σ	d	q	s
d	$S/2$	0	0
q	0	$S/2$	0
s	0	0	Q

(12.5)

where Q is the kinetic energy stored in the rotor, due to the application of the outside mechanical torque f

$$\mathbf{Q} = -\int f_s dx^s. \quad (12.6)$$

Hence analogously to the electrical stress-energy tensor (10.7) the mechanical stress-energy tensor can be defined as

$\overset{m}{T}_{\alpha}^{\beta}=$

$\alpha \backslash \beta$	d	q	s	t
d				
q		$\overset{m}{\boldsymbol{\pi}}$		$\overset{m}{\mathbf{G}}$
s				
t		$\overset{m}{\mathfrak{S}}$		$\overset{m}{\mathbf{W}}$

(12.7)

$\overset{m}{T}_{\alpha}^{\beta}=$

$\alpha \backslash \beta$	d	q	s	t
d	$S/2$	0	0	0
q	0	$S/2$	0	0
s	0	0	Q	mv
t	0	0	Sv	$-mv^2/2$

(12.8)

Here also $$T_{\alpha}^{\alpha}=0. \tag{12.9}$$

The *mechanical impulse-energy equation* can be written analogously to Eq. (11.1)

$$\delta \overset{m}{T}_{\alpha}^{\beta}/\partial x^{\beta}=\overset{m}{p}_{\alpha}. \tag{12.10}$$

(1) Let the free index $\alpha=s$ (geometrical axis). Assuming the dummy indices $\beta=t$

$$\delta \overset{m}{T}_{s}^{t}/\partial x^{t}=\delta \mathbf{G}/\partial t=\partial \mathbf{G}/\partial t=mdv/dt, \tag{12.11}$$

representing the inertial torque.

Assuming the dummy index $\beta=s$ (geometrical)

$$\delta \overset{m}{T}_{s}^{\pi}/\partial x^{\pi}=\delta \overset{m}{T}_{s}^{s}/\partial x^{s}=-\delta Q/\partial x^{s}=-f, \tag{12.12}$$

representing the negative of the applied outside torque.

Hence the equation of torque of Lagrange (1.5) may be written as

$$\delta \overset{m}{\mathbf{G}}/dt+\text{Abs. Div. }\overset{m}{\boldsymbol{\pi}}=\mathbf{p}, \tag{12.13}$$

where $\mathbf{p}$ is the torque developed by the machine.

(2) Let the free index $\alpha=t$ (time axis). Assuming the dummy indices $\beta=t$,

$$\delta \overset{m}{T}_{t}^{t}/\partial x^{t}=\delta \overset{m}{\mathbf{W}}/\partial x^{t}=-d\overset{m}{W}/dt, \tag{12.14}$$

representing the time rate of change of stored kinetic energy.

Assuming the dummy index $\beta=s$

$$\delta \overset{m}{T}_{t}^{s}/\partial x^{s}=\partial \overset{m}{T}_{t}^{s}/\partial x^{s}=\partial S_{v}^{s}/\partial x^{s}$$
$$=v^{s}\partial S/\partial x^{s}=pv^{s}, \tag{12.15}$$

representing that part of the electrical power input that changes into mechanical power.

The mechanical equation of power (2.3) may be written as

$$\delta \overset{m}{\mathbf{W}}/dt+\text{Abs. Div. }\overset{m}{\mathfrak{S}}=-\mathbf{f}v. \tag{12.16}$$

Hence $\overset{m}{p}_{\alpha}$ represents the positive electrical torque and the negative of the mechanical power input.

XIII. The Combined Impulse-Energy Equation

The electrical and mechanical stress-energy tensors (10.7) and (12.7) can be combined into one tensor as

$$\overset{em}{T}_{\alpha}^{\beta}=\overset{e}{T}_{\alpha}^{\beta}+\overset{m}{T}_{\alpha}^{\beta}, \tag{13.1}$$

$\overset{em}{T}_{\alpha}^{\beta}=$

$\alpha \backslash \beta$	π	t
π	$\overset{e}{\boldsymbol{\pi}}+\overset{m}{\boldsymbol{\pi}}$	$\overset{e}{\mathbf{G}}+\overset{m}{\mathbf{G}}$
t	$\overset{e}{\mathfrak{S}}+\overset{m}{\mathfrak{S}}$	$\overset{e}{\mathbf{W}}+\overset{m}{\mathbf{W}}$

(13.2)

$\overset{em}{T}_{\alpha}^{\beta}=$

$\alpha \backslash \beta$	π	t
π	$\overset{em}{\boldsymbol{\pi}}$	$\overset{em}{\mathbf{G}}$
t	$\overset{em}{\mathfrak{S}}$	$\overset{em}{\mathbf{W}}$

(13.3)

It satisfies the relation

$$\overset{em}{T}_{\alpha}{}^{\alpha}=0. \tag{13.4}$$

The combined impulse-energy equation is

$$\delta\overset{em}{T}_{\alpha}{}^{\beta}/\partial x^{\beta}=\overset{em}{p}_{\alpha}. \tag{13.5}$$

It represents the complete equation of torque (1.5) and the complete equation of power (2.1) of the dynamical equations of Lagrange.

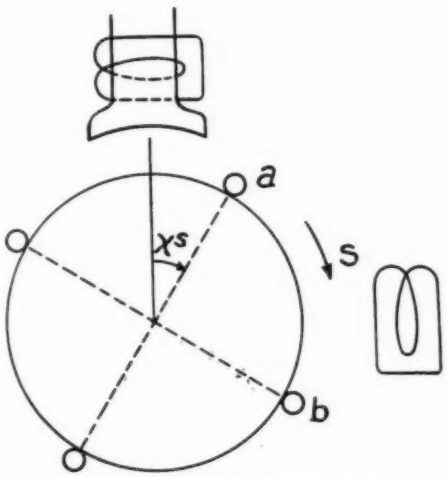


FIG. 3. Holonomic reference axes.

If the free index $\alpha=s$ (geometrical) the equation of torque (1.5) becomes

$$\delta\overset{em}{\mathbf{G}}/\partial t+\text{Abs. Div. }\overset{em}{\boldsymbol{\pi}}=\overset{em}{\mathbf{p}}_{s}=0. \tag{13.6}$$

If the free index $\alpha=t$ (time) the equation of power (2.1) becomes

$$\delta\overset{em}{\mathbf{W}}/\partial t+\text{Abs. Div. }\overset{em}{\mathfrak{S}}=\overset{em}{\mathbf{p}}_{t}, \tag{13.7}$$

$$\overset{em}{p}_{t}=-e_{a}i^{a}-fv. \tag{13.8}$$

The equation of power (13.7) may also be written as

$$\delta\overset{em}{\mathbf{W}}/\partial t=\overset{em}{\mathbf{p}}_{t}, \tag{13.9}$$

since the internal energy-flow, Abs. Div. $\overset{em}{\mathfrak{S}}$ cancels.

XIV. The Equations of Lagrange in Terms of Field Quantities

If the *field tensor* $F_{\alpha\beta}$, Eq. (5.14) is used, the equation of motion of Lagrange (1.1) can be written (if the free index γ does not assume $t=$ time, but the dummy indices do) as

$$a_{\pi\alpha}\frac{d^2x^{\alpha}}{dt^2}+[\alpha\beta,\,\pi]\frac{dx^{\alpha}}{dt}\frac{dx^{\beta}}{dt}+F_{\pi\alpha}\frac{dx^{\alpha}}{dt}=f_{\pi}, \tag{14.1}$$

where now f_{π} is the applied mechanical torque. The applied voltage E_{π} is included in $F_{\pi t}x^{t}$.

If the free indices represent only electrical indices, then the *equation of voltage*, (1.4) *representing the motion of the electric charges, is*

$$a_{m\alpha}\frac{d^2x^{\alpha}}{dt^2}+[\alpha\beta,\,m]\frac{dx^{\alpha}}{dt}\frac{dx^{\beta}}{dt}+F_{m\alpha}\frac{dx^{\alpha}}{dt}=0. \tag{14.2}$$

This equation now is in the same form as it appears in the unified field theory, with the right-hand side equal to zero.[5]

This same equation may also be written in terms of the *vector potential* φm

$$\frac{d\varphi m}{dt}+F_{mu}\frac{dx^{u}}{dt}=0, \tag{14.3}$$

$$\frac{d\varphi m}{dt}-2S_{mu}{}^{\gamma}\varphi_{\gamma}\frac{dx^{u}}{dt}=e_{m}. \tag{14.4}$$

In terms of the combined stress-energy tensor (13.2) the remaining dynamical equations, namely the equation of torque (1.5) and the equation of power (2.1) are represented by the combined impulse-energy equation (13.5)

$$\delta\overset{em}{T}_{\alpha}{}^{\beta}/\partial x^{\beta}+\overset{em}{p}_{\alpha}=0. \tag{14.5}$$

XV. Holonomic Reference Frames

If the reference axes are rigidly connected to the moving conductors (Fig. 3) then two cases are possible:

(1) The vector-potential is defined along the moving axes as

$$\varphi_{\pi}=a_{\pi\sigma}i^{\sigma}. \tag{15.1}$$

[5] Einstein and Mayer, Einheitliche Theory von Gravitation und Elektrizität. Sitzungsberichte der Preussischen Akademie der Wissenshaften. Phys.-Math. Klasse **25**, (1931), Eq. (28a).

In this case the current density wave s^α, the field tensor $H^{\alpha\beta}$, etc., have no existence in general and outside of the potential equation

$$e_m = d\varphi_m/dt \tag{15.2}$$

most of the field equations become nonexistent.

(2) First a quasi-holonomic reference frame is introduced along the d and the q axes and the vector-potential φ_a is defined along these axes. Then *the vector-potential along the holonomic axes a and b is defined in terms of its components along the quasi-holonomic reference frame*, namely in terms of φ_d and φ_q as

$$\varphi_\pi = \begin{array}{c|cc|cc} \pi & a & & b & s & t \\ \hline & \varphi_d \cos x^a + \varphi_q \sin x^a & & -\varphi_d \sin x^a + \varphi_q \cos x^a & 0 & 0 \end{array} \tag{15.3}$$

$$\varphi_\pi = \begin{array}{c|cc|cc} \pi & d & & q & s & t \\ \hline & \varphi_a & & \varphi_b & 0 & 0 \end{array} \tag{15.4}$$

With this definition of φ_π *the usual holonomic forms of the tensor field equations can be used*, with $S_{\alpha\beta}{}^\gamma = 0$. All tensors s^α, $H^{\alpha\beta}$ etc., have an existence and have the same absolute value as when expressed along the quasi-holonomic reference frame. For instance

$$F_{\alpha\beta} = \delta\varphi_\alpha/\partial x^\beta - \delta\varphi_\beta/\partial x^\alpha = \partial\varphi_\alpha/\partial x^\beta - \partial\varphi_\beta/\partial x^\alpha \tag{15.5}$$

$$= F_{\alpha\beta} = \begin{array}{c|cccc} \alpha \backslash \beta & a & b & s & t \\ \hline a & 0 & 0 & \varphi_b & e_a \\ b & 0 & 0 & -\varphi_a & e_b \\ s & -\varphi_b & \varphi_a & 0 & 0 \\ t & -e_a & -e_b & 0 & 0 \end{array} \tag{15.6}$$

since

$$\partial\varphi_a/\partial x^t = d\varphi_a/dt = e_a, \tag{15.7}$$

$$\partial\varphi_a/\partial x^s = -\varphi_d \sin x^s + \varphi_q \cos x^s = \varphi_b. \tag{15.8}$$

All successful electrical engineering theories follow this second method of attack.

In case of the *stator* all axes on it are holonomic. However no quasi-holonomic axes can be assumed on the stator and so no s^α, $F_{\alpha\beta}$, $H^{\alpha\beta}$, p_α etc. can be defined for it in general, except for the special cases of moving magnetic paths and moving reference axes.

Time Constants of Incandescent Lamps

W. E. Forsythe, Miss M. A. Easley and D. D. Hinman
Incandescent Lamp Department, General Electric Company, Nela Park, Cleveland, Ohio

Candescence and nigrescence curves are shown for a number of tungsten lamps, vacuum and gas-filled, ranging in size from the automobile lamps up to the 30-volt, 30-ampere lamp and the 1000-watt photoflood lamp. Data are given for similar filaments mounted in vacuum lamps and in lamps filled with argon, nitrogen, hydrogen, and helium. The filament of a vacuum lamp reaches 90 percent brightness on heating in time $t=0.060I^{0.54}$; 10 percent brightness on cooling in time $t=0.022I^{0.70}$; for gas-filled lamps, $t=0.15I^{0.70}$ and $t=0.070I^{0.70}$, respectively. (I is the normal operating current of the lamp.) The percent of total variation, P, in brightness during a cycle for a vacuum lamp on 60-cycle a.c. is $P=34D^{-1.06}$, on 25-cycle a.c., $P=28D^{-1.37}$; for a gas-filled lamp, $P=75D^{-0.89}$ and $P=63D^{-0.91}$, respectively. (D is the filament diameter in mils.)

WHEN a constant voltage is applied to an incandescent lamp, the current suddenly rises much higher than its normal value. The value of this sudden rush of current is determined by the cold resistance of the filament as compared both to its final hot resistance, which is about fifteen to seventeen times the cold resistance, and to any resistance between the lamp and the source of power. The initial current for the smaller wattage 115-volt tungsten lamps may be several times the final value, but the difference for the larger lamps (about 500 watt and larger) is much less because the resistance of the external circuit may be of the same order as, or even larger than, the cold resistance of the lamp. If the circuit external to the lamp filament is such that it does not change resistance appreciably due to the final current through the lamp, the temperature of the tungsten filament of a vacuum lamp will not rise above its operating temperature due to the sudden rise of current. Whether the initial rush of current will heat the filament of a gas-filled lamp above its normal operating conditions depends upon several factors such as gas loss, number of supports, lamp voltage, and lamp wattage. However, our experiments have shown that the light output of the 115-volt gas-filled lamps does not initially reach a value higher than its final value. The length of time that the current is higher than its final value is the time it takes the filament to reach its maximum temperature providing there is no overshooting in the temperature of the filament. This time is determined by the amount of heat energy that is stored up in the filament and the amount of energy that is radiated during this time.

The amount of energy stored in the filament depends upon the temperature of the filament, its physical dimensions, and its specific heat. The time for the regular rate of input of energy to the filament to supply this amount of energy has been called the modular[1] time of the lamp (see Table I). These modular times are not the times necessary for constant applied wattage to raise the filament to temperatures of normal operation because as soon as the filament is raised to a temperature higher than its surroundings, it begins to lose energy by the methods already mentioned and these losses increase the amount of energy necessary and thus the time required to reach the normal temperatures. Under conditions of constant applied voltage, the filament will heat up somewhat more rapidly as the energy is supplied at a greater rate. For constant current, however, the filament will heat somewhat more slowly than at constant wattage since constant current represents a slower supply of energy due to the fact that the filament has a lower resistance when cold than when hot.

The times required, after the application of voltage, for the filaments of lamps of a number of different sizes to reach such a temperature that they radiate 90 percent of their light output at normal temperature as well as the times required, after the applied voltage is turned off, for the filaments to cool to such a temperature that they give only 10 percent of their normal light were measured by two methods and the results obtained are shown in Figs. 1 and 2 and in Tables I, II, and III with some other operating characteristics of the lamps. The lamps were heated by current from a storage battery and the external

[1] Forsythe and Adams, G. E. Review **39**, 497 (1936).

circuit resistance including that of the storage battery was about 0.4 ohms.

The first set of measurements was made with the rotating flashometer[2] which gave a photographic trace from which, using appropriate density calibrations, the relative amounts of light radiated at different times could be determined.

[2] Forsythe and Easley, R. S. I. **5**, 216 (1934).

For the second set of measurements, a photometer consisting of a photoelectric tube with a selected filter such as is used in the regular testing of the light output of incandescent lamps was used as the light-evaluating instrument. The output of the photoelectric tube, after amplification by means of a specially constructed amplifying system, was recorded by an oscillograph.

TABLE I. *Time[1] characteristics of some 115-volt incandescent lamps.*

Lamp	L/W	Temp. °K	Modular Time	Time of Heating to 90% Brightness	Time of Cooling to 10% Brightness	Time of Cooling to 400% R	Time of Cooling to 200% R	Range of Variation in Light Output (Percent) 60 cycles	25 cycles
Vacuum Lamps									
6 w	6.9	2400	0.04	0.039	0.012			74[2]	103
10 w	8.2	2465	.06	.057	.018			58	92
40 w	10.8	2570	.14	.128	.058	1.00		14	40
50 w sign	10.8[3]	2590	.15	.130	.051			13	36
Gas-filled Lamps									
40 w	11.1	2710	.06	.065	.026	.48	1.3	27	58
60 w	12.9	2770	.09	.102	.042	.52		18	44
60 w CC	13.9	2765	.12	.110	.051	.63			
100 w	15.6	2845	.13	.125	.059	.72	2.3	13	33
200 w	17.5	2880	.20	.220	.094	1.50		8	22
300 w	19.0	2915	.27	.270	.125	1.65		6	17
500 w	20.3	2955	.36	.380	.190	3.2	12.0	4.5	13
Photoflood Lamps									
No. 1, 250 w	34.	3490	.13	.125	.075	1.25		8.1[2]	
No. 2, 500 w	31.6	3400	.20	.225	.120			5.1[2]	
No. 4, 1000 w	32.0	3410	.91	.400	.210			3.0[2]	
Flood and Spot Lighting Lamps									
250 w Floodlt.	15.	2900	.27	.260	.115			6.6[2]	
400 w Spotlt.	20.	3055	.22	.340	.160			6.0[2]	
250 w Spotlt.	17.7	2930	.22		.110			7.0[2]	
30 V-30 A Movie	26.5	3290	1.10	1.15	.48			1.3	
Hydrogen-filled 115-volt Lamps									
250 w		2910		.24	.055				
400 w		3085		.37	.072				
1000 w monoplane		3160		.31	.16				
biplane				.23	.14				
Helium-filled Lamps									
90 V-136 w		2910		.25	.055				
85 V-192 w		3085		.26	.065				
100 V-860 w		3160		.39	.18				

[1] Time in seconds. [2] From curve. [3] Blue bulb, clear bulb efficiency.

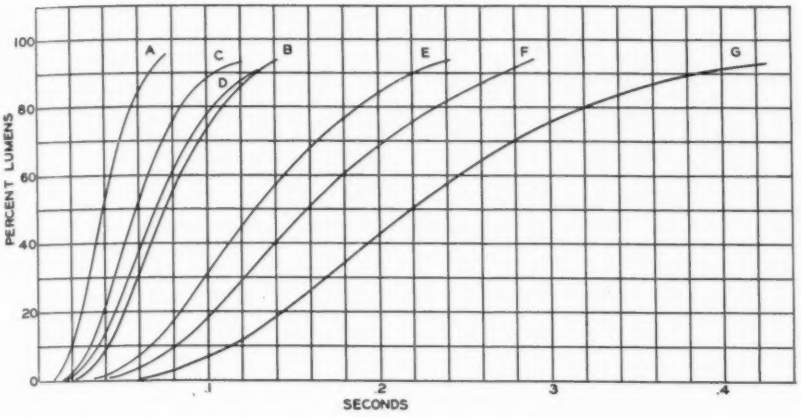


FIG. 1. Starting times of various 115-volt tungsten filament incandescent lamps. Curve A, 40 w gas-filled; curve B, 40 w vacuum; curve C, 60 w gas-filled; curve D, 100 w gas-filled; curve E, 200 w gas-filled; curve F, 300 w gas-filled; curve G, 500 w gas-filled.

When an ordinary 115-volt incandescent lamp is cooled to a point where it gives off only 10 percent of its normal light output, the filament is at a temperature between 2000 and 2300°K. When it has heated to a point where it will give 90 percent of its normal light output, the filament is only about 1 percent low in temperature, that is, about 30°K below its normal operating temperature. The times required after turning on the current for the filaments of different 115-volt vacuum lamps to reach such a temperature that they gave off 90 percent of their final light output were found to be represented by the following equation:

$$t = 0.060 I^{0.54}$$

(t = time and I = normal operation current).

The times of heating to 90 percent brightness for the gas-filled, 115-volt lamps having a life of from about 500 to 1000 hours were found to be represented by the following equation:

$$t = 0.15 I^{0.70}.$$

The times required after turning off the current for the different sized 115-volt vacuum lamps to cool to such a temperature that they give 10 percent of their normal light output are given by the following equation:

$$t = 0.022 I^{0.70}.$$

The corresponding cooling times for the 115-volt gas-filled lamps are found to follow the equation:

$$t = 0.070 I^{0.70}.$$

TABLE II. *Times of heating and cooling for a special filament lamp filled with various gases.*

Lamp	Voltage for Temperature of 3000°K	Watts	Starting Time to 60% Brightness	Starting Time to 90% Brightness	Stopping Time to 10% Brightness
Hydrogen-filled	115	350	0.075	0.340	0.072
Helium-filled	85	190	.125	.260	.066
Nitrogen	73	141	.135	.210	.086
Vacuum	68	124	.140	.210	.096

TABLE III. *Time constants of some gas-filled automobile lamps.*

Lamp	Volts	Current	Candlepower	Temp. °K	Time of Heating to 90% Brightness	Time of Cooling to 10% Brightness
No. 63	6.85	0.53	2.9	2820	0.071	0.028
55	6.5	.40	1.9	2715	.070	.026
2330	6.2	4.13	32.	2970	.29	.141
2320 *L*						
No. 1 fil.	6.2	4.67	35.8	2955	.36	.15
No. 2 fil.	6.4	2.98	21.8	2905	.24	.12

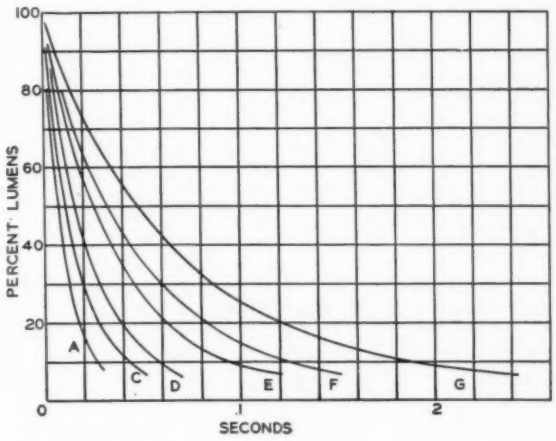


FIG. 2. Stopping times for various 115-volt gas-filled tungsten filament lamps. Curve A, 40 w; curve C, 60 w; curve D, 100 w (40 w vaccum about same as this); curve E, 200 w; curve F, 300 w; curve G, 500 w.

The times of heating and cooling are shown in Fig. 3 plotted on logarithmic scale against the normal current in amperes. Calculations of the time characteristics using figures on time

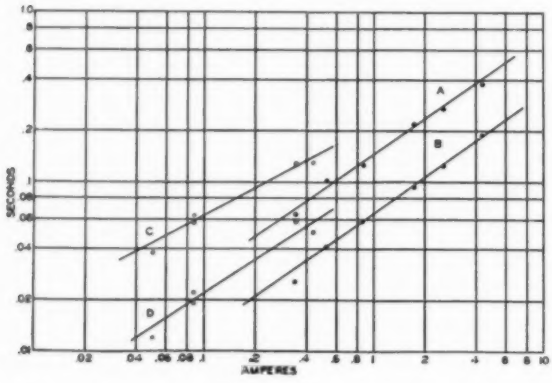


FIG. 3. Time characteristics of some 115-volt tungsten filament incandescent lamps; curve A, starting time to 90 percent brightness for 40 w–500 w gas-filled tungsten lamps; curve B, times of cooling to 10 percent brightness; curve C, times of heating to 90 percent brightness for 6 w–50 w vacuum tungsten lamps; curve D, cooling times to 10 percent brightness for 6 w–50 w vacuum tungsten lamps.

characteristics given by Jones and Langmuir[3] agree well with our data on vacuum lamps.

These equations are similar to those found by Sanford and Waldram[4] with slightly different values of the constants since the times given in their paper are slightly different from the time characteristics given here.

Since it is difficult to measure the time of cooling with the photo-tube photometer or the rotating flashometer for values of brightness less than 5–10 percent of the maximum value, another method had to be used for measuring times for cooling to lower values. This was done by use of a pendulum device to open the lamp circuit at a definite time before making instantaneous electrical connections to a potentiometer circuit for measuring the resistance of the filament. By repeated trials the resistance could be quite accurately measured during the instantaneous connections at these different times after the current through the lamp was broken and from the resistance the temperature of the filament could be determined if necessary. By this method the times that it took a number of lamps to cool to 400 percent and to 200 percent of their cold resistance were measured and are shown in Table I. The tungsten lamp filament for 200 percent cold resistance is at a temperature of about 250°C and at 625°C for 400 percent resistance. The filament of a tungsten lamp is just visible in a dark room when the filament is heated to about 300 percent of its cold resistance, that is, to a temperature of about 450°C which means 4.6 volts applied to the 115-volt, 500-watt gas-filled lamp, 9.4 volts applied to the 40-watt gas-filled lamp, and 3.0 volts applied to the 25-watt vacuum lamp.

By calculation from the data in Tables I and II and other characteristics of tungsten, it was found that it would require about a 60 percent increase in the times given in the table for the different lamps to cool from 10 percent brightness to such a temperature that they would give 1 percent of the normal light output. Also calculations showed that an additional 15 percent in time would be required to heat the different filaments from 90 percent to 95 percent brightness.

The starting time for some large lamps with resistance in series with the lamp was measured. As would be expected, the resistance slowed down the starting of the lamp. The time required for some of the 115-volt photoflood lamps to heat from a temperature (about 2500°K) corre-

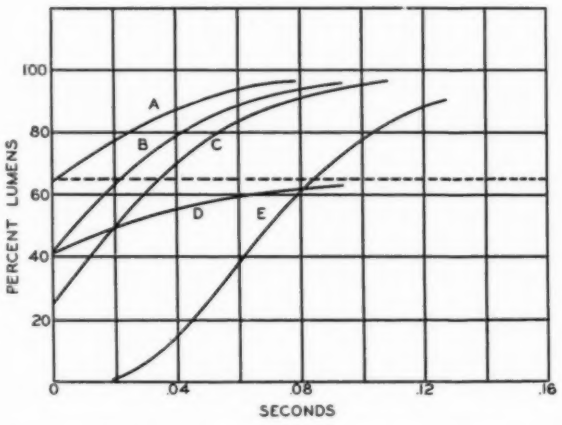


FIG. 4. Heating times of 100 w, 115 v gas-filled tungsten lamps for various conditions. Curve E, 115 volts applied to cold filament; curve A, 15 volts additional applied to lamp already operating on 100 volts; curve B, 25 volts additional applied to lamp already operating on 90 volts; curve C, 35 volts additional applied to lamp already operating on 80 volts; curve D, 10 volts applied to lamp already operating on 90 volts.

[3] Jones and Langmuir, G. E. Review **30**, 354 (1927).
[4] Illum. Eng. (London) **28**, 109 (1935).

sponding to 50 volts to 90 percent maximum brightness was also measured. With the lamp at the brightness corresponding to 50 volts applied

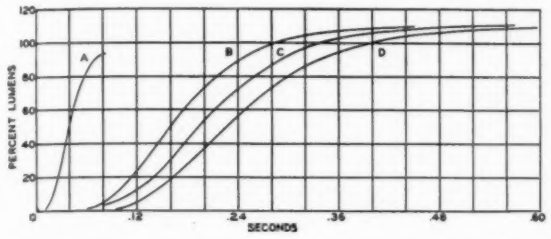


FIG. 5. Starting times for some automobile lamps. Curve A, Mazda No. 63; curve B, Mazda 2320L—2nd filament; curve C, Mazda 2330; curve D, Mazda 2320L—1st filament.

to the lamp, the starting curve obtained was practically parallel to the original starting curve. The No. 4 (1000-watt) photoflood lamp reached about 10 percent of the maximum brightness when operated on 50 volts. When 115 volts was applied to the lamp, 10 percent brightness was reached in about 0.10 sec., and 90 percent brightness in 0.40 sec. When the lamp was operating on 50 volts and suddenly an additional 65 volts was applied, 90 percent brightness was reached in 0.30 sec.

The time of heating of the regular 100-watt gas-filled tungsten lamp was measured from 0 to 115 volts, from 80 to 115 volts, from 90 to 115 volts, from 100 to 115 volts, and from 80 to 100 volts. The results are shown in Fig. 4.

The starting and stopping times were measured for some lamps filled with other gases than argon. Such lamps should heat faster for gases of higher heat loss since, for the corresponding wattage, the filament is smaller in diameter, but on the other hand the gas slows down the time of heating. The net result was that in general hydrogen or helium-filled lamps of corresponding wattage heated to 90 percent brightness more slowly than the regular lamps, but cooled more quickly. The heating curve for some of the hydrogen filled lamps showed what appeared to be a break when the lamp was heated to about 50 to 60 percent brightness and from there on the filament heated at a much slower rate. This may be due in part to the breaking down of the hydrogen molecule, but the relation between initial input, filament dimension and gas losses (bulb temperature) is largely responsible for this change in rate of heating of hydrogen filled lamps.

Several lamps made up using the filament and mount for a special 115-volt, 350-watt hydrogen-filled lamp were filled with nitrogen, helium and with hydrogen and others were exhausted to a good vacuum. Each lamp was operated at the voltage required to produce a filament temperature of about 3000°K. The times of starting and stopping of these lamps were measured and the results are given in Table II.

The starting curves did not show any evidence of overshooting for any of these lamps. The hydrogen and helium-filled lamps start to heat up faster, but at above 60 to 70 percent brightness the rate becomes slower and they require a longer time to reach 90 percent brightness than do the other lamps. This is caused by the relative power input and the gas losses (bulb temperature). The times of heating and cooling for this type of lamp filled with the argon-nitrogen mixture used for commercial lamps were about the same as for nitrogen-filled lamps.

In Table III and Fig. 5 are given the time constants of certain automobile lamps. Many of these lamps, in particular those requiring large current, initially reached a brightness higher than that of their normal operating temperature. This was due to the relationship between the filament resistance and that of the lamp circuit external to it. The leads inside the lamp are heated during operation due to their own resistance and also to the heat conducted to them from the filament. The result of this is that the current does not drop back to the operating value until the leads reach their final temperature, which helps cause the overshooting. For the same reason the 30-volt, 30-ampere movie lamp overshoots in brightness about 5 percent and is about 2 percent brighter than normal ten seconds after it is turned on.

When an incandescent lamp is operated on a.c. the temperature of the filament and thus its light output changes as the current passes from its maximum value to zero. The cyclic variation in the light output for 60-cycle a.c. is too rapid to be directly observed, but it can be measured by either of the methods used to measure the heating and cooling times of the different lamps.

The cyclic variation in the light output for 60-cycle and 25-cycle a.c. was measured for a number of various wattage lamps ranging in size from the 115-volt, 40-watt lamp to the 30-volt, 30-ampere movie lamp, and the logarithm of the percent, P, of total variation in the light output was plotted against the logarithm of the filament diameter, D, and a linear relation found as shown in Fig. 6. This cyclic variation of the light output of the different gas-filled lamps operated on 60-cycle a.c. voltage is given by the following equation:

$$P = 34D^{-1.06}$$

and the corresponding variation for the vacuum lamp is represented by the following equation:

$$P = 28D^{-1.37}.$$

When operated on 25 cycles, the following equations give the results for gas-filled lamps and for vacuum lamps, respectively.

$$P = 75D^{-0.89},$$
$$P = 63D^{-0.91}.$$

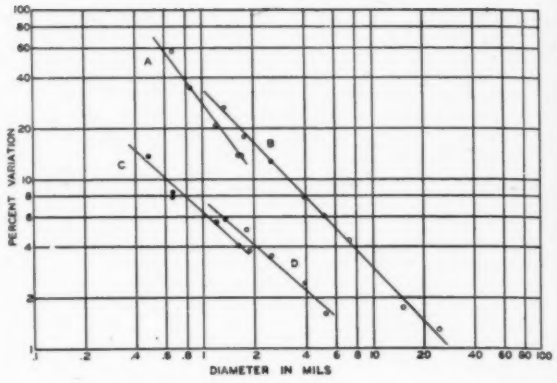


FIG. 6. Percent variation in brightness as a function of filament diameter of various tungsten 115-volt lamps when operating on alternating current. Curve A, vacuum lamps on 60-cycle, a.c.; curve B, gas-filled lamps on 60-cycle, a.c.; curve C, vacuum lamps on 25-cycle, a.c.; curve D, gas-filled lamp on 25-cycle, a.c. (The percent scale for curves C and D should be multiplied by 10.)

These results show why at times one can notice the flicker of tungsten filament incandescent lamps on 25-cycle a.c. while the flicker is not noticeable on 60-cycle a.c.